

ECONOMIC AND FINANCIAL CRISES

**A NEW MACROECONOMIC
ANALYSIS**

**ALVARO CENCINI
AND SERGIO ROSSI**



Economic and Financial Crises

Also by Alvaro Cencini

CAPITOLI DI TEORIA MONETARIA

ELEMENTI DI MACROECONOMIA MONETARIA

MACROECONOMIC FOUNDATIONS OF MACROECONOMICS

MONETARY MACROECONOMICS: A New Approach

MONETARY THEORY, NATIONAL AND INTERNATIONAL

MONEY, INCOME AND TIME

TIME AND THE MACROECONOMIC ANALYSIS OF INCOME

EXTERNAL DEBT SERVICING: A Vicious Circle (*with Bernard Schmitt*)

LA PENSÉE DE KARL MARX: Critique et synthèse (*with Bernard Schmitt*)

PER UN CONTRIBUTO ALLO SVILUPPO DEL MICROCREDITO (*with Marco Borghi*)

CONTRIBUTI DI ANALISI ECONOMICA (*co-edited with Mauro Baranzini*)

INFLATION AND UNEMPLOYMENT: Contributions to a New Macroeconomic Approach (*co-edited with Mauro Baranzini*)

Also by Sergio Rossi

LA MONETA EUROPEA: UTOPIA O REALTÀ? L'emissione dell'ecu nel rispetto delle sovranità nazionali

MACROECONOMIE MONETAIRE: Théories et politiques

MODALITES D'INSTITUTION ET DE FONCTIONNEMENT D'UNE BANQUE CENTRALE SUPRANATIONALE: Le cas de la Banque Centrale Européenne

MONEY AND INFLATION: A New Macroeconomic Analysis

MONEY AND PAYMENTS IN THEORY AND PRACTICE

MACRO E MICROECONOMIA: Teoria e applicazioni (*with Mauro Baranzini and Giandemetrio Marangoni*)

MODERN MONETARY MACROECONOMICS: A New Paradigm for Economic Policy (*co-edited with Claude Gnos*)

MODERN THEORIES OF MONEY: The Nature and Role of Money in Capitalist Economies (*co-edited with Louis-Philippe Rochon*)

MONETARY AND EXCHANGE RATE SYSTEMS: A Global View of Financial Crises (*co-edited with Louis-Philippe Rochon*)

THE ENCYCLOPEDIA OF CENTRAL BANKING (*co-edited with Louis-Philippe Rochon*)

THE POLITICAL ECONOMY OF MONETARY CIRCUITS: Tradition and Change in Post-Keynesian Economics (*co-edited with Jean-François Ponsot*)

Economic and Financial Crises

A New Macroeconomic Analysis

Alvaro Cencini

*Professor of Economics and Chair of Monetary Economics, University of Lugano,
Switzerland*

Sergio Rossi

*Professor of Economics and Chair of Macroeconomics and Monetary Economics,
University of Fribourg, Switzerland*

palgrave
macmillan



© Alvaro Cencini and Sergio Rossi 2015

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2015 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–1–137–46189–6

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Contents

<i>List of Tables and Figures</i>	vii
<i>Acknowledgements</i>	viii
Introduction	1
Part I Modern Principles of Monetary Macroeconomics	
1 The Monetary Macroeconomics of Modern Economic Systems	15
2 The Macroeconomic Laws of Monetary Production Economies	38
Part II Business Cycle and Crisis Theories: A Fundamental Critique	
3 Business Cycles versus Boom-and-Bust Cycles	59
4 From Monetarism to the New Classical Synthesis	83
5 From Keynes to Post-Keynesian Economics	106
6 Economic Crises and Their Relationship to Global Supply and Global Demand	129
Part III The Monetary Macroeconomics of Crises	
7 Capital Accumulation and Economic Crises	151
8 Interest, Rate of Interest, and Crises	184
9 The International Dimension of Financial Crises	204
10 Reforming Domestic Payment Systems	226
11 Reforming the International Monetary System	241

<i>Bibliography</i>	260
<i>Author Index</i>	271
<i>Subject Index</i>	273

Tables and Figures

Tables

1.1	The opening of a line of credit	20
1.2	The payment of wages	25
1.3	The purchase of produced output	36
1.4	The result of the final purchase of output	37
4.1	The use of money-marble as unit of payment	88
9.1	The equality between A's real imports and its real exports	224
10.1	The result of the payment of wages through the two departments of a bank	229
10.2	The result of a bank loan	231
10.3	The result of a residential mortgage loan	232
10.4	The entry of profit in the second department	235
10.5	The transfer of profit to the third department	235
10.6	The investment of profit	236
10.7	The entry of redistributed and invested profit	237

Figures

1.1	The result of the payment of wages	26
5.1	The IS–LM diagram	110
5.2	The relationship between S and I	113
5.3	The identity between L and M	115
7.1	The process of fixed capital formation	173
7.2	The production of amortization goods	176
7.3	The expenditure of the income formed in the production of amortization goods	177
7.4	The investment of profit	178
7.5	The production of amortization goods and its consequences	179
9.1	The double charge of the payment of in	213

Acknowledgements

Bernard Schmitt and his quantum macroeconomics have been a constant, privileged source of inspiration for us in the preparation of this volume. Remembering and honouring his groundbreaking work, his humanity, and his tireless support, we pay grateful tribute to him, a mentor and a friend, who passed away on 26 March 2014. We are also grateful to our undergraduate, graduate, and postgraduate students for their thought-provoking questions and comments. Helga Wild and Niklas Damiris have generously helped us by revising the English manuscript and also by providing useful suggestions and comments. Andrea Carrera, research and teaching assistant at the University of Lugano, Switzerland, has skilfully plotted all figures and tables and completed the bibliography. Our warmest thanks go to all of them.

Introduction

Confronted with the most serious economic and financial crises since the 1930s, economists should feel the need to question their approach to economic analysis as well as the conceptual background of mainstream economics. Their mathematically sophisticated models, whether neoclassical, new classical, Keynesian, or New Keynesian, have clearly proved to be incapable of avoiding, let alone explaining, the devastating crises that are hampering our economies both nationally and internationally. The reason for this failure lies in their poor understanding of the logical laws governing our economic systems based, as they are, on bank money.

As surprising as this might appear, inflation, involuntary unemployment, sovereign debts, financial bubbles, and global economic recessions are all negative effects of a single original cause: the erroneous conception of bank money. The import of this claim can only be evaluated once it is recognized that the emission of bank money is what characterizes all our economies. In the absence of banks, only pre-capitalist economies would exist. Without bank money, neither monetary nor financial intermediations would be possible, and the terms 'inflation', 'deflation', 'financial bubble', and 'sovereign debt' would be meaningless. Yet, this is not to say that these pathologies are the unavoidable consequences of the discovery of double-entry book-keeping and the creation of banks. Banks make it possible to build an economic system based on the accumulation of capital. Whether such a system is an orderly or a disorderly one depends on whether it conforms or not to the nature of money, income, and capital. Bank money in itself is not the cause of any pathology; rather, it is the way money is kept distinct from or is mixed up with income and capital that determines whether the result is pathological or not.

The importance of bank money has been clearly perceived by many great economists of the past. From Smith's distinction between nominal and real money to Keynes's emphasis on the principle of double-entry bookkeeping, one can trace a line that culminates with the work of Schmitt and his unique definition of bank money. What distinguishes quantum macroeconomic analysis from mainstream economics is the definition of money and the relevance attached to it. According to quantum monetary analysis, money is a numerical form issued by banks as an asset–liability and associated with physical output through the payment of wages. Its great relevance derives from the fact that it enables the numerical expression of wages and makes it possible to express real goods numerically, in terms of wage units. Economics would not exist as a science if its object of enquiry could not be measured. As the Classics knew well, the determination of a unique and invariable standard of value plays a crucial role in the building of economics as a science. Yet, the search for such a standard is unavoidably hopeless so long as economic value is considered as a dimension. It was Walras who first recognized that economic value is essentially a numerical relationship and not a substance as the Classics believed. Unfortunately, Walras did not follow up on his intuition and was unable to find the numerical standard of value implied by it. It is with Keynes's introduction of the wage units that a solution has appeared. But it is only once money is identified as an asset–liability that Keynes's wage units acquire their full significance in economic analysis.

If it is the case (and who could deny it?) that our economic systems are monetary, and that money is bank money, then it is also a fact that if economic laws exist at all, they must be related to money and to the way the latter is associated with production. It is at this stage that another important distinction appears between quantum macroeconomics and mainstream economics. According to the latter, economic laws are strictly related to economic agents' behaviour (the all too famous law of supply and demand is the clearest example of such microeconomic laws), whereas according to the former, economic laws are objectively defined by the identities deriving from the presence of bank money as well as from its flow nature and its role as a means of payment. Now, the choice between these two antithetical approaches is not a matter of preference. Either macroeconomic identities exist in reality or they do not. If they do, they define the strictest possible relationship between their terms, a relationship that holds good whatever the behaviour of economic agents and that is independent of any set of norms imposed on monetary and financial institutions. It is only a rigorous analysis of

our monetary economies of production that can establish whether economic identities are a matter of fact or not. And bank money is the necessary starting point of such an analysis. In this volume, we consider the various steps that, starting from the analysis of bank money, allow one to discover the logical laws of our monetary economies and explain the nature of the pathologies as deriving from the lack of conformity of the present systems of national and international payments with these laws.

Since its origin, economic analysis has been conceived either as a tool to understand a system fundamentally based on relative exchange and on economic agents' behaviour or as an instrument to discover the structural logical laws that economic agents have to comply with. The distinction between these two kinds of analyses has not always been clear-cut, and in the last decades differences have become blurred. On the whole, apart from rare exceptions, it might be claimed that economists are now unanimous in believing that, conceptually, nothing fundamental has yet to be discovered in their domain. The great majority of them are utterly convinced that their efforts must tend towards the construction of mathematical models that are able to reproduce economic reality more and more accurately, in an attempt to provide ever more reliable economic forecasts. Substantially, they view economics as an empirical 'science' akin to meteorology and consider economic crises as unfortunate events caused by unexpected external shocks. If this view were correct, the role of economists would be to make the best models to anticipate the impact of such external shocks in order to mitigate against their possible negative effects.

Now, this pragmatic approach to economics has proved unsatisfactory in practice and can additionally be shown to be wrong conceptually. Its conceptual poverty is revealed immediately as soon as one raises the question concerning the nature of money and its 'integration' with physical output. To consider money as a net asset or, even worse, as a commodity is a misleading conception and a dangerous metaphysical claim. The confusion between money and income is widespread and would not be worth dwelling on if it had no serious consequences. Alas, this is not the case. The correct understanding of the nature of money and of the payment that transforms it into income is crucial, because it leads straightforwardly to the discovery of the first logical law of macroeconomics: the identity between any macroeconomic supply and its demand.

Neoclassical and Keynesian attempts to develop mathematical models capable of reproducing the real-world dynamics fail because economics

is not a branch of mathematics, and this is so because macroeconomic laws are simply logical identities. For too long, economics has been trapped in an axiomatic theoretical framework characteristic of the methodology in the field of mathematics. General equilibrium models of a Walrasian or non-Walrasian type are clear examples of this kind of approach, and Keynesian general disequilibrium models of various types, although distinct from the former, are not substantially different for the simple reason that they too are founded on a series of conditional equalities whose terms balance only for the equilibrium value of some specific variables. Essentially, economic models are far too ambitious. Their aim to reproduce as complex a reality as that of capitalist economies is way beyond the capabilities of mathematical modelling. But even if they managed to account technically for all the conceivable decisions taken by economic agents, for their combinations, and for all the possible external shocks, the mathematical and the statistical approaches would still have to be rejected. The reason for this rejection lies in the nature of the laws at the core of any economic system based on the accumulation of capital.

The difference between mainstream economics and the quantum macroeconomic analysis advocated in this volume could not be clearer: whereas mainstream economics, whether of a neoclassical or a Keynesian mould, is built around the notion of equilibrium between supply and demand, Schmitt's macroeconomic analysis is founded on the necessary equality between these two terms when referred to the same production. Whereas mainstream economists look for the origin of economic and financial crises by investigating the relationship between external shocks and economic agents' behaviour, quantum macroeconomists look for it at the structural level, investigating the actual workings of the system of national and international payments. Whereas for mainstream economics inflation and unemployment are to some extent unavoidable and the only thing that can be done, or at least attempted, is to keep them under control, for the advocates of quantum monetary macroeconomics these are pathological conditions that can be eradicated through a reform of the present structure of national payments.

The differences between the analysis developed by Schmitt, as introduced again in this volume, and the traditional approach based on the microeconomic foundations of macroeconomics are not just confined to the study of national economies of production, but they also concern the enquiry into the causes of international disorders related to the international economy of exchange. As a matter of fact, mainstream

economics has little to say about the pathological state of the international economic system. Since Keynes's and Schumacher's proposals for a reform of the system of international payments, mainstream economists have provided no further insight into this matter, and their explanation of international disorders, such as the sovereign debt crisis or the creation and exponential growth of a financial bubble, remains therefore highly unsatisfactory. They recognize that the present 'system' is in reality a 'non-system' of international payments, but they do not explain why it is so, or what has to be done to transform it into a sound structural arrangement. Once again, what is missing in their analysis is a correct understanding of the nature of money and of the way international payments have to be carried out. What they do not know is that money is essentially a flow (and not a stock in motion), and that this becomes more apparent at the international level, where real goods produced nationally are traded among countries.

The net-asset definition of money is a deleterious source of serious mistakes and a barrier against the correct understanding of what is needed to provide the world with a sound system of national payments. The most relevant improvements in this area have been accomplished by practising bankers rather than by academic economists. National banking systems owe their present structure to bankers, and it is only fair to recognize that it is thanks to their initiative (and their advisors' recommendations) that countries are homogeneous monetary areas, which means that a unique currency exists within each sovereign country even though each single commercial bank issues its own spontaneous acknowledgement of debt (money). Monetary homogeneity is the result of a system of real-time gross final settlements created by banks and managed by the central bank. Surprisingly enough, neither economists nor bankers have so far realized that, unless a structurally similar system is created internationally, national currencies are bound to remain heterogeneous and payments between countries pathological.

Quantum monetary macroeconomics is based on a thorough investigation of the nature of bank money and provides a new insight into the character of international payments. In particular, it shows that countries' sovereign debts are entirely pathological, as their very existence is due to the absence of a proper system of international payments. The gravity of the sovereign debt crisis is all too real to insist on the relevance of Schmitt's investigation. What is immediately unclear in the analysis of mainstream economists is their meaning or understanding of the term 'sovereign debt'; so it is not surprising to observe that sovereign debt is often wrongly identified as being the debt incurred by the State.

The public debt, however, is not co-extensive with the debt of a country defined as the set of its residents, which owes its existence as much to private as to public debts incurred abroad. This lack of a correct definition of sovereign debt is the unavoidable result of misunderstanding the way countries are involved in the external payments carried out by their residents (State included). Schmitt's analysis makes it clear that sovereign debts are of a macroeconomic nature and therefore cannot be imputed on the (mis)behaviour of economic agents. Once again, it is the non-system of international payments that is identified as the cause of this monetary pathology.

The discovery of a pathological duplication affecting the payment of a country's net interest on debt as well as the formation of a country's external debt, its sovereign debt, is Schmitt's last legacy and a key result of quantum macroeconomic analysis. Its relevance for the understanding of financial crises is great, and its outcome is particularly significant because of the reform it calls for. To impute the sovereign debt crisis to the excess of borrowing that countries incur in order to finance their surplus expenditures amounts to maintaining that the balance-of-payments principle, which establishes the necessary equality between each country's total purchases or imports (commercial and financial) and its total sales or exports (commercial and financial), is systematically disregarded by the present non-system of international payments. As a matter of fact, the disregard for the balance-of-payments identity is what characterizes the pathological state of the actual system. This does not mean that, as any other logical identity, the necessary equality between each country's total sales and purchases can be put in jeopardy by the behaviour of economic agents, in particular by their decision to increase their foreign expenditures. Logical identities cannot be transformed into conditions of equilibrium and cannot be submitted to the goodwill of economic agents. However, so long as these laws are not fully understood and complied with, a discrepancy will always arise between them and the way payments are carried out by banks. When the identity concerning payments among countries is not complied with, then the pathology that emerges from this lack of conformity generates a country's sovereign debt. Plainly stated, this amounts to claiming that sovereign debts should not exist, that their very formation is of a pathological nature, and that they can and must be avoided through a reform that allows for the implementation of a system of international payments consistent with the balance-of-payments identity.

The aim of economic analysis is to explain the real world of economics and to provide a solution to the economic and financial crises

that are currently plaguing it. Mainstream economics has failed on both fronts, mainly because it has, erroneously, assumed that the logical foundations of macroeconomics are microeconomic. On the contrary, quantum macroeconomic analysis shows that macroeconomics has its own – that is, macroeconomic – foundations, which consist in a set of logical laws that form the analytical framework for the orderly working of the national and international systems of payments. The reforms presented in this volume are those developed by Schmitt between 1984 and 2014 and pertain to the structural changes necessary to avoid national and international economic and financial crises respectively, as well as monetary and financial disorders originating from international transactions. Their common feature is the aim to provide a system of payments respectful of the numerical and flow nature of money.

Indeed, the choice between mainstream and quantum economics is about two radically different ways of coping with economic and financial crises. The microeconomic approach chosen by mainstream economics considers crises the unavoidable result of a system in constant search of an ever-fleeting equilibrium. In such a framework, the least economists are expected to do is to reduce disruptive fluctuations to a minimum, being aware that unexpected shocks are always lurking and ready to prove them wrong at any time. In short, we would just have to learn to live with inflation, unemployment, and sovereign debt in the hope to be able to limit their amplitude as much as we can. On the other hand, quantum macroeconomics provides a novel way out of this predicament by proposing a structural bookkeeping reform of both the national and international systems of payments, able to eradicate the causes of the pathology affecting our economies. Let us be very clear in this respect. The reforms needed to make the national and international systems of payments consistent with the macroeconomic laws deriving from the logical distinction between money, income, and capital are not solutions to the problems concerning what and how to produce; how much to pay different categories of workers; how to redistribute income; and what role to attribute to the State, to trade unions, to lobbies, and so on. What these reforms make possible is only the passage from a disorderly to an orderly system of payments, both nationally and internationally. In a reformed economic system, which we could name a regime of post-capitalism if we call capitalism the actual pathological system, inflation, involuntary unemployment, and sovereign debt will no longer be possible. However, this key shift still only involves providing a sound structure that guarantees the ‘neutrality’ of money. People will still have to decide on what economic policy to implement

to support the ideal society of their choice. The role of economists is thus precise and circumscribed: to specify the structural reforms required to avoid the pathological working of our economic systems. This is the aim of the present volume, which unfolds as follows.

Chapter 1 asks a number of fundamental questions that are still to be answered properly by the economics profession at large. What is money and how do banks issue it? Is money endowed with a positive value since its creation, or does it acquire its purchasing power, and how? Is it necessary to distinguish between money and income? The chapter explains that the nature of money remains a mystery and needs to be investigated starting from its bookkeeping origin. A rigorous analysis based on double-entry bookkeeping shows that money is intrinsically valueless and can only derive its purchasing power from production. This new macroeconomic analysis of money leads to dismiss the old-fashioned idea that by issuing money banks originate credit, in terms of a loan granted to the economy and financed by banks themselves.

Chapter 2 elaborates on the positive analysis of monetary macroeconomics and discusses the macroeconomic laws of monetary production economies. Following Walras's contribution, and despite Keynes's suggestions, today it is generally accepted that economic laws are mainly behavioural and founded on microeconomics. This is instrumental for the use and abuse of mathematics, and equations of various kinds are considered as the best tools available to represent the real world of economics. The aim of this chapter is to re-establish the logical priority of identities and to show that identities are in fact the foundation of macroeconomic analysis. For that purpose, the second chapter starts from a reappraisal of Say's law, goes on to give a reinterpretation of Keynes's identity between global demand and global supply, and ends with Schmitt's law of necessary equality between each agent's sales and purchases.

Chapter 3 analyses the search for a theory of crises where economic disturbances are considered as endogenous events inherent in the workings of our economic systems. Such an approach is essentially macroeconomic and aims to determine the laws supposedly intrinsic to capitalism. The chapter also addresses business cycle theories that aim to show that crises are periodical events due to economic fluctuations. Whether in the form of business cycle theories emphasizing the role played by trade, money, and credit, or in the form of real business cycle theories, the models proposed in this framework have in common their microeconomic structure. They identify thereby the causes of economic disorders in exogenous shocks imputable to agents'

(mis)behaviour. This approach also underpins Minsky's boom-and-bust cycle theories of financial crises.

Chapter 4 shows that both monetarism and the new classical synthesis fail to provide a satisfactory analysis of the working of our economic systems and of the way disorders may arise in them. It argues that the so-called equation of exchange on which monetarism rests is tautological, and that the concept of the 'quantity of money' is completely at odds with the true nature of bank money. New classical economics fares no better. This approach provides new general equilibrium models explaining business cycles consistently with the microeconomic approach typical of neoclassical analysis. The rational expectations hypothesis plays a crucial role in these models, which attempt to identify the causes of economic crises in irregular external shocks and imperfect information. Whether in the form of monetarist or new classical models, in them money continues to have little or no impact and is still considered as a commodity or identified with a financial asset.

Chapter 5 investigates Keynesian, New Keynesian, and post-Keynesian economics to verify if they succeed in reaching a better understanding of the origin of crises than their neoclassical counterpart. It shows that Keynesian economists of all schools fail to reach this goal, despite their emphasis on the role played by monetary factors. Keynesian and New Keynesian economists indeed aim at finding adequate microeconomic foundations for their models and thus have abandoned any attempt to search for the macroeconomic foundations of macroeconomic analysis. The emphasis that some post-Keynesian economists put on the role of money and banks in a monetary production economy may lead one to conclude that their approach is much closer to the message conveyed by Keynes's own analysis than that of Keynesian and New Keynesian economists. However, post-Keynesian economists have completely lost sight of the conceptual distinction between money and credit.

Chapter 6 investigates economic crises and their relationship to global supply and global demand. Starting from Say's law and Keynes's logical identity between Y and $C + I$, the chapter addresses the problem of whether or not the insurgence of an economic crisis entails the rejection of them. Indeed, the possibility of reconciling a situation of disequilibrium with the identity of global supply and global demand seems inexistent. However, quantum macroeconomics provides logical evidence that the identity between global supply and global demand is at the heart of economics. This can only mean that, eventually, economic crises will have to be explained without denying this identity. The

chapter argues that economic crises can be explained by simultaneously respecting the identity between global supply and global demand and by allowing for a numerical difference between them.

Chapter 7 focuses on capital accumulation to detect if the latter can lead to economic crises. Capital is indeed one of the central concepts of economics. However, there is still no consensus among economists on how to define it. Well-known economists such as Ricardo, Marx, Walras, Böhm-Bawerk, and Keynes have addressed this question and, with the notable exception of Walras, have reached the conclusion that capital cannot be considered as a direct source of economic value. However, both Ricardo and Böhm-Bawerk emphasize the role played by time in enabling capital to be an indirect source of economic value. Keynes's analysis of capital is another important contribution to a correct understanding of this concept and encapsulates all the deepest insights of his predecessors concerning the role of saving and time. By introducing these elements into a theoretical framework where the presence of both money and banks is essential, Keynes opens the way to the modern macroeconomic analysis of capital.

Chapter 8 elaborates on the analysis presented in the preceding chapter, focusing on interest and interest rates. Capital is indeed formed through the investment of profit and defines a macroeconomic saving: this derives from Keynes's identities between global supply, Y , and global demand, $C + I$, and between S and I . A correct analysis of interest has to respect these identities and explain how it is possible to derive a positive macroeconomic value from capital given that labour is the sole macroeconomic factor of production. The chapter shows that capital accumulation reduces the rate of profit within the economic system taken as a whole. By narrowing the gap between the rate of profit and the market rate of interest, this creates the conditions for an economic crisis. As the macroeconomic rate of profit gets closer to the market rate of interest, investment must be reduced, and this has a negative effect on employment. The economic crisis is then worsened by the financial crises induced by a growing pathological capital and the speculative transactions it feeds.

Chapter 9 focuses on the analysis of international transactions and their impact on financial crises. It shows that what is wrong with the system of international payments is actually the way it works. In particular, it introduces two strictly related analyses that have led to the discovery of a pathological duplication of countries' debts. The first concerns the problem of indebted countries' external debt servicing and shows how the payment of net interest on a country's external debt has

actually a total cost of twice the amount of the interest due to foreign creditors. The second analysis explains that the pathological duplication induced by transnational payments entails the very formation of countries' external debts and not merely the payment of net interest on these debts.

Based on the arguments in its preceding chapter, Chapter 10 presents a structural monetary reform of domestic payment systems. As the global financial crisis that erupted in 2008 has made it plain, banks are money as well as credit providers. In fact, as Ricardo explained, the emission of money and the provision of credit can be carried out by two separate bodies, without the slightest loss of advantage. Indeed, this separation is a structural factor of financial stability, because it allows avoiding that banks issue empty money in purely financial transactions that do not generate new income within the economic system as a whole. The chapter elaborates on this and explains also the importance of introducing a third department in banks' accounting, which has to account for those profits that are invested in the purchase of capital goods and which therefore should not be available as bank deposits to finance lending operations on any kind of market. This will avert capital over-accumulation and the resulting macroeconomic disorders.

The last chapter of this volume, Chapter 11, deals with the necessary reform of the international monetary system. This chapter presents the reform elaborated by Schmitt on the basis of his analysis of the pathological formation of countries' sovereign debt. The advantage of this reform is that it can be implemented by any single country irrespective of what is done by the rest of the world and without causing any harm to it. Thanks to this reform, a country would be able to avoid the pathological duplication of its external debt, and its government's budget would earn the domestic income lost today because of the net expenditures carried out by its residents. Besides showing that any single country can protect itself against the monetary and financial disorders caused by the present non-system of international payments, the chapter also shows that the passage to an orderly system of international payments is possible. In particular, the euro area could implement easily enough a reform preventing its member countries to suffer from the serious drawbacks caused by the actual lack of finality of their external payments and from their sovereign debt crisis.

On the whole, the positive and normative analysis presented in this volume shows that it is not only possible but also urgent to transform economic analysis into a socially useful and powerful tool for human development in a framework where systemic economic and financial

crises cannot occur. This should be a welcome contribution to provide a really scientific status to the 'dismal science' whereby economics and economists can provide a set of policy proposals aimed at the common good within an orderly working economic system, nationally as well as internationally.

Part I

Modern Principles of Monetary Macroeconomics

This page intentionally left blank

1

The Monetary Macroeconomics of Modern Economic Systems

What is money and how do banks issue it? How is it associated with physical output? Is it endowed with a positive value since its very creation, or does it acquire its purchasing power, and how? Is it necessary to distinguish between money proper and money income? If yes, what is their logical relationship? These are some of the questions dealt with in this chapter, where we endeavour to show that the specific nature of bank money remains, in part, a mystery and needs to be investigated starting from its bookkeeping origin. Too often identified with a commodity or an asset, money is mainly perceived as a stock that can circulate more or less rapidly within the economy and whose cost has a direct impact on production. Is this definition consistent with the way money enters those payments that banks carry out? A rigorous analysis based on double-entry bookkeeping shows that this is not the case, because money is intrinsically valueless and can only derive its value or purchasing power from production. The classical distinction between nominal and real money finds a new *raison d'être* in the distinction between money and income, where the latter is the result of a transaction through which money (a simple numerical form) integrates produced output as its real content. This new macroeconomic analysis of money and income leads to a fundamental dismissal of the old-fashioned idea that money creation is nothing less than a credit creation, or in other words that, by issuing money, banks originate credit, that is, a loan granted to the economy and financed by banks themselves. In fact, banks act as monetary as well as financial intermediaries, and credit is never financed through money creation.

Let us proceed step by step along a path that will lead us to discover the principles of modern monetary macroeconomics.

About money

Economists almost invariably begin their monetary writings by giving a wide and usually empirically based definition of money, which may correspond to the dictum that money is what money does, or may go as far as to identify money with a commodity, an asset, or a veil. Even the author of *A Treatise on Money* is no exception, for he starts his famous book by defining money through its functions.

Money itself, namely that by delivery of which debt contracts and price contracts are *discharged*, and in the shape of which a store of general purchasing power is *held*, derives its character from its relationship to the money of account, since the debts and prices must first have been expressed in terms of the latter.

(Keynes 1930/1971: 3)

The weakness of this approach is that it assumes that a conceptual definition may be made to coincide, a priori, with a nominal or with an axiomatic definition. Nominal definitions are arbitrary and say nothing about the nature of what is being thereby defined. The choice of a word to appose to an object is a clear example of nominal definition. Whatever word we choose to name a given object or concept, our understanding of its nature does not progress at all. To the extent that it does not increase our knowledge, an axiomatic definition is no more useful than a nominal one. Moreover, these two kinds of definition are quite different. Axiomatic definitions, in fact, are not arbitrary. As universally established principles, they are the result of a process of understanding, which transforms them into self-evident statements only *a posteriori*. A true axiom is a principle arrived at through analytical discovery. The fact that the Earth rotates around the Sun may be taken as an axiom today, but was certainly not considered as such before Copernicus. Finally, correct conceptual definitions are bound to become axioms. When this happens, they can be taken for granted and introduced as axiomatic definitions from the outset. What cannot be done is to assume the axiomatic character of a definition before having rigorously established it. A conceptual definition is indeed the arrival point of a process of understanding, and not its point of departure logically. As far as monetary analysis is concerned, we cannot start by axiomatically defining money, because the definition of money must be the end result of our conceptual enquiry.

The different forms that money is supposed to take on, and the different ways in which it is made to operate, are symptomatic of the lack of

a clear understanding of what money really is. As claimed by Mishkin (2004: 44), '[e]conomists define *money* (also referred to as the *money supply*) as anything that is generally accepted in payment for goods or services or in the repayment of debts'. On top of being aprioristic, this definition is far too vague to be meaningful. As a matter of fact, anything can be accepted in repayment of debts, from banknotes and coins to checks, deposit certificates, securities, or even goods and services. The only apposite conclusion in this regard is therefore that 'there is no single, precise definition of money or the money supply, even for economists' (ibid.: 44).

In 1930, Keynes entitled Book I of his *Treatise* 'The Nature of Money' and Book II, 'The Value of Money'. This is the logical succession that has to be followed if we are to avoid taking for granted any aprioristic conception of money. Claiming, as done by Mankiw (2007: 77), that '*money* is the stock of assets that can be readily used to make transactions' is twice misleading, since it is at the same time too broad and too specific a definition. It is too broad because it encompasses every kind of assets, and too specific because it assumes that money is an asset. In reality, one cannot take for granted that money is an asset, even though historical observation seems to corroborate it. While it is undisputable that gold, silver, and other materials have been used in the past to represent money, it is no less certain that money cannot be identified with any of these materials. To describe the physical aspect of a coin, a banknote, or any other object is no contribution to understanding what money's nature is.

In their explanation of money, economists almost invariably start from a brief analysis of the different forms taken up by what has historically been chosen to play the role of a standard. The passage from commodity money to bank money has marked a process of increasing dematerialization, which clearly shows how erroneous it would be to keep identifying money with its physical supports. Money is not gold, nor silver, nor an electric impulse; convertibility has long been abandoned; and central banks have given up the official link between money and their gold reserves. The recent evolution of payment systems towards the electronic use of book-entry money has clarified the terms of the problem, thus making it easier for researchers to avoid confusion. However, the main difficulty remains, namely that of explaining the nature of an entity that, although dematerialized, pertains to the realm of economics. Money is essentially a conceptual entity, which is not identifiable with any material or any object whatsoever, but is nevertheless strictly associated with or integrated to produced goods and services. Abstraction is not an easy way to follow. However, it is a necessary step,

which has to be taken without falling into the trap of abandoning any reference to the real world. Money exists in the real world, but it cannot be physically defined or identified; this is why it has so imperfectly been understood so far. This explains also why it has generally been confused with a series of physical objects to which it was erroneously identified. Our problem is thus clearly stated: to determine the specificity of money irrespective of the device used to play its role.

The nature of nominal money

Let us start from the first function usually attributed to money, namely that of a unit of account. Economists are not unanimous when defining this function of money. For many of them, money acts as unit of account in that it is a measure or a standard of value. The 'role of money is to provide a *unit of account*; that is, it is used to measure value in the economy' (Mishkin 2004: 46). For others, money is a numerical unit used to express economic magnitudes such as prices, debt, income, capital, and so on. 'As a *unit of account*, money provides the terms in which prices are quoted and debts are recorded' (Mankiw 2007: 77). Now, to define money as a standard of value from the outset is logically unacceptable, since it would amount to suppose that money has a positive value, and that its value is of the same kind as (and therefore comparable with) that of the goods and services it is meant to measure. Such a procedure must be rejected, for the simple reason that, as intuited by Keynes (1930/1971), the nature of money must be explained first, before asking whether money has a value of its own or not. The numerical conception of money avoids this criticism and is far more promising to understand the working of our economic systems.

A unit of account is, first of all, a numerical standard used to count or to enumerate. Strictly speaking, it consists of numbers alone, and is therefore deprived of any intrinsic value. Numbers can be used to count a collection of homogeneous goods in order to ascertain their arithmetical sum. This is not, however, what a unit of account is supposed to do in economics, where produced goods and services are far from being homogeneous. In this framework, numbers are given the task of representing the means through which physically heterogeneous objects may be given a common numerical form. To introduce money as a unit of account has no other meaning than introducing numbers in the realm of economics.

The earlier statement may sound strange. Numbers are just here for us to use; why should we need money to introduce them into economics? The answer relates to our previous remark: goods and services

are physically heterogeneous and cannot be summed up unless they are made commensurable. The existence of numbers is not enough to solve this problem. What is still needed is a way to integrate numbers and goods, which requires numbers to be made available through a specific economic operation.

Numbers have been part of economics since the introduction of money. However, it is only with the appearance of an exhaustive monetary system and through the generalized use of money that the integration of numbers has been achieved. The easiest way to understand the numerical nature of money is therefore that of analysing what allows banks to provide the economy with a unit of account.

The idea that banks create money is not new. Keynes (1930/1971) devotes the first section of his chapter on bank money to the 'creation' of bank money, and claims that 'the bank may create a claim against itself in favour of a borrower, in return for his promise of subsequent reimbursement' (Keynes 1930/1971: 21). It is by spontaneously issuing their own acknowledgement of debt that banks create money, and it is worth observing here that, correctly, Keynes does not confine the creation of money to the central bank. Any bank can issue money by creating a claim against itself. The question that has to be clarified at this stage is how it is possible for a bank to get spontaneously indebted, and how it is that, through its spontaneous acknowledgment of debt, numbers are introduced into economics.

The discovery of double-entry bookkeeping, which took place in thirteenth century's Italy, is the event that marks the origin of banks. Double-entry bookkeeping was itself made possible by the previous (seventh century) discovery, attributed to the Indian mathematician Brahmagupta, of negative numbers, which, in its turn, is closely related to a new conception of the number zero. For a long time, zero was not considered a number in its own right: it was merely conceived as a symbol for an empty space in the system of numeration; it represented the absence of anything, the void, or the 'nothing'. Since Brahmagupta, zero is known to have a definite numerical value of its own. More precisely, zero is the first number of the series of positive integers, the number that precedes one and that separates positive from negative numbers. Double-entry bookkeeping is an application of this distinction, which allows arriving at zero by adding negative to positive numbers.

The first meaning of bank money creation is the possibility given to banks by double-entry bookkeeping to issue simultaneously $+x$ and $-x$ units of money. Indeed, this is the only acceptable way of envisaging creation without recurring to metaphysics. Banks are human institutions,

Table 1.1 The opening of a line of credit

Bank B			
Assets		Liabilities	
Credit on client A	<i>x</i> m.u.	Debit to client A	<i>x</i> m.u.

and as such they cannot create anything positive. The risk of running into metaphysics is avoided if what is created is at the same time positive and negative. This is what happens when banks get spontaneously indebted to a client who accepts to get indebted to them, that is, when a bank agrees to open a credit line to the benefit of one of its client, A. The double-entry representation of this contract is shown in Table 1.1.

By entering A on the liabilities side of its ledger, bank B acknowledges its willingness to get indebted to its client A insofar as the latter is willing to become its debtor. Overdraft facilities pertain to this kind of transactions. They simply mean that the bank is prepared to carry out a payment on behalf of its client, and that its client will have to reimburse it. No payment has yet occurred, and the amount of money actually created is zero. However, the operation is not void and meaningless. The simultaneous creation of a positive and a negative amount of money to the benefit of A has no direct consequence for B and for A, but sets the conceptual and the practical framework for a payment system to exist. The signal given by banks to the economy is that they can provide the numerical vehicle required to convey a payment.

At this point a new question arises. If money is but a numerical vehicle with no value, and if banks by themselves can only issue zero units of money, how can any payment be carried out? The answer to this question leads us straightforward to consider the problem of money's value.

The value of money

Apparently, the easiest way to attribute a positive value to money would be to define it as an asset. However, as we know, this is not a suitable solution, because bank money is almost completely dematerialized, and, even more important, because this definition cannot be introduced as an axiomatic assumption. An alternative solution is offered by the chartalist view that money in general may be accepted as a means of payment, as one of its components is State money. As a matter of fact, economists unanimously agree that the passage to inconvertibility has

generalized the use of fiat money, that is, ‘paper currency decreed by governments as legal tender (meaning that legally it must be accepted as payment for debts)’ (Mishkin 2004: 48). A government decree would thus be at the origin of money’s general acceptance in payments. This is so much so as the government may impose different tax liabilities to the population and declare that it accepts State money as a means for discharging these liabilities. According to Smith (1776: 328), ‘[a] prince, who should enact that a certain proportion of his taxes should be paid in a paper money of a certain kind, might thereby give a certain value to this paper money’ (see also Mitchell Innes 1913: 398–9). Hence, the State can induce all taxpayers to accept these pieces of paper as money, because any non-bank agents know for sure that everyone who has to pay taxes will accept them in turn (Tobin and Golub 1998: 27).

Attractive as the chartalist view might seem at first, it takes us no nearer to a solution of the problem, for it rests on the arbitrary assumption that the government has the supernatural power to create a currency with a positive redeeming power. ‘[Fiat money] is manufactured by the government from thin air’ (Tobin 1965: 676). As already emphasized by von Mises (1912/1981), the chartalist theory of money does not provide any explanation of either prices or money’s value, and must be rejected as entirely useless. ‘The state theory is not a *bad* monetary theory; it is not a monetary theory at all’ (von Mises 1912/1981: 510).

A way out of this conundrum seems to be that of referring to social, general acceptance, that is, of maintaining that money’s redeeming power derives from ‘its general acceptability in the discharge of public and private transactions’ (Tobin 1965: 676). So, we are told that money has a positive value because we all accept it as a means of payment discharging debts. Money’s value would derive from social convention, in a similar way as in language the meaning of words is determined universally because everybody agrees about it.

As long as everyone continues to accept the paper bills in exchange, they will have value and serve as money. Thus, the system of commodity money evolves into a system of fiat money. Notice that in the end the use of money in exchange is a social convention: everyone values fiat money because they expect everyone else to value it.

(Mankiw 2007: 80)

Social, general agreement emerges as the cause of money’s value on top of government’s enforcement of paper currency as unique means of tax payments. Yet, a question arises almost spontaneously: for what reason

should individuals agree on accepting a mere acknowledgment of debt in exchange for their goods and services? Bank money being deprived of any intrinsic value, its general acceptance should derive from the generalized belief that everybody is prepared to accept it as if it had a positive value. To put it bluntly, this amounts to saying that if we all accept to be dupe, no one of us will eventually be fooled. Needless to say, this is an odd way of making ends meet. Indeed, the whole argument is fallacious, and winds up in a vicious circle, money's general acceptance being founded on the belief that everybody will accept it. Reality is much more straight and pragmatic: it is because money has a positive value (whose origin we still have to explain) that it is generally accepted, and not the other way around.

In most general terms, money's value is identified with money's purchasing power. What makes an individual to accept money in exchange for her/his goods and services is the certainty that the sum received will allow her/him to purchase other items of an equal value. The pertinent question to ask, therefore, is where does money's purchasing power come from. To explain how money acquires a positive purchasing power is to explain how we can pass from nominal to real money, to put it in the language of classical economists, or, using today's terminology, from money to income. 'We mean by the purchasing power of money the power of money to buy the goods and services on the purchase of which for purposes of consumption a given community of individuals expend their money income' (Keynes 1930/1971: 48). Let us address this issue in the next section.

The logical and factual distinction between money and income

As we noted in the previous section, double-entry bookkeeping empowers banks to spontaneously incur a debt to the economy. Since this debt is balanced at once by an equivalent credit with respect to the same agent, the opening of a credit line has no direct consequences either on the creation of money, or on income. In this first phase, banks simply show that they are able and willing to provide a numerical unit of account to their clients, and to convey payments on their behalf. What has to be clearly understood is that, up to this point, no money has been issued, and no income has been formed. Hence, banks can, using double-entry bookkeeping, 'vehiculate' payments, but they are by no means capable to finance them so far. No shortcut is permitted here. Logically, we must start from *tabula rasa*, and we cannot suppose the

pre-existence of positive bank deposits of any sort. The formation of income has to be explained starting from a situation in which banks have no deposit, hence no income to lend.

The problem we are faced with is how to determine the payment that banks can carry out starting from a zero income. At first sight, it would seem that logic leaves no room for a positive answer. If payments pertained all to the same category and implied the final purchase of goods and services, this would be it. Indeed, any final purchase implies the expenditure of a positive income, of which there is no trace at the beginning of our thought experiment. The absence of any pre-existing income is not the only requirement imposed to the analysis, which forbids also the pre-existence of produced goods and services. It follows that banks cannot finance any final purchase both because no income is available in the system so far, and because there are no goods to be purchased at this stage. Our task is thereby drastically simplified: to find a payment that does not require the expenditure of a positive income and does not define the purchase of any produced good or service.

As argued by Keynes (1930/1971: 111), income is directly related to production. 'We propose to mean identically the same thing by the three expressions: (1) *the community's money income*; (2) *the earnings of the factors of production*; and (3) *the cost of production*'. The payment we are looking for is therefore a payment that banks can carry out on behalf of firms, and which does not amount to the purchase of raw materials, energy, machinery, or any other item of this kind. Now, only one candidate is eligible for this: the payment of wages. This clearly means that, unlike what Marx (1867/1976) wrongly assumed, the payment of wages does not define the purchase of a particular commodity called labour-power. It also means that the payment of wages does not define the final purchase of produced output by firms or banks. When paying wages to their workers, firms would be purchasing the result of their labour activity only if wages were paid out of a positive income. If this were indeed the case, income would be deemed to remain totally unexplained. If a positive income exists – and no one doubts it – it is because, as a result of the payment of wages, firms do not become the final owners of produced output.

Most authors advocating a macroeconomic approach to economics have emphasized the peculiarity of human work. Among them we find the exponents of classical theory, as well as Keynes and some of his followers. According to these authors, labour is not a commodity, but rather the source of all of them. In particular, classical economists identify labour with the source of economic value, and

Keynes (1936/1946: 213–4) considers labour as the sole macroeconomic factor of production. ‘It is preferable to regard labour, including, of course, the personal services of the entrepreneur and his assistants, as the sole factor of production, operating in a given environment of technique, natural resources, capital equipment and effective demand’. As a matter of fact, it is not only preferable but also logically compulsory to consider labour as essentially different from all other (microeconomic) factors of production. This is so because labour does not result from any process of production. Land and capital are themselves the result of production, whereas labour is not. In terms of inputs and outputs this may be expressed by saying that labour is nothing but an input, while land and capital are outputs before becoming (microeconomic) inputs: they must first be produced if they are to be used in the production of some other outputs.

The final proof that labour is the sole (macroeconomic) factor of production has been provided by quantum economic analysis, and rests on the fact that workers alone can be credited with the result of a payment that does not require the expenditure of a pre-existing income.

Indeed, let us recall that money as such is the spontaneous acknowledgment of debt issued by banks, and that the object of this debt is a payment. What the beneficiary of the bank’s acknowledgement of debt obtains is the promise that, when required, a payment will be carried out on her/his behalf. Now, banks carry out their payments by using double-entry bookkeeping, that is, by debiting the payer and crediting the payee into their ledgers. The payment we are concerned with here regards production, and is meant to have the double task of accounting for the positive creation of money, and the formation of income. In this regard, Schmitt (1998–99a: 52, our translation) holds the view that the only apposite payment is the payment of wages, as ‘*individuals* only are credited and debited. Neither land, nor capital can be the subject of credits and debits. [...] Once again, it is inconceivable that a capital be credited or debited or “credited-debited”. Only the “human factor” is suitable for that’.

By opening a credit line to firms, F, banks declare their availability to carry out any payments up to an agreed amount. If a positive income were already deposited with banks, their initial payment on behalf of F could be made to the benefit of agents selling any kind of goods, for example raw materials or machinery. As no production has yet occurred, no pre-existing deposits are available, and income amounts to zero. Under these circumstances, the initial payment of banks cannot consist in the purchase of any positive asset, either real or financial. What

Table 1.2 The payment of wages

Bank B			
Assets		Liabilities	
Credit on firm F	x m.u.	Wage earners' deposit	x m.u.

is, then, the nature of the payment of wages? How is it that banks can pay F's workers even though they have no positive income at their disposal? To answer these questions we have to analyse in some detail the payment of wages as entered in a bank's ledger.

Let us start from the double-entry bookkeeping representation of the three-pole transaction defining the payment of wages (Table 1.2).

The three poles concerned by the payment of wages are the bank, B, which acts as a simple intermediary, firm F, on behalf of which the payment is carried out, and workers, W, who are credited in the bank's ledger. What can be immediately observed is that the payment represented in Table 1.2 defines a complete and self-consistent transaction. From a double-entry point of view, the payment does not require any complement, since what is entered on the assets side of B's balance sheet is perfectly matched by what is entered on its liabilities side. It is also immediately clear that W's credit to the bank is not compensated by any debit incurred by W. Workers are net creditors, which means that, since they have been paid out of zero income, their payment is at the origin of a positive income. To be sure, the source of money's value is labour, and the payment of wages gives it its numerical form. The meaning of what is entered on the liabilities side of B's ledger is univocal: workers are the owners of the new income deposited with B. The bank owes W the amount W lend to it by depositing their newly formed income.

The formation of income is closely related to production. Double-entry bookkeeping alone could not account for it. The passage from money to income takes place at the very moment production is monetized, that is, when physical goods and services are given a common numerical form through their association with money. The nature of money is thus double: it is both a flow and an immaterial form. These joint aspects of the same entity have a very short existence, which coincides with the instant wages are paid. As a flow, money comes alive in every payment, and disappears as soon as the payment has occurred. As a numerical form, money carries out its task at the moment physical goods and services are given their numerical expression. In the payment

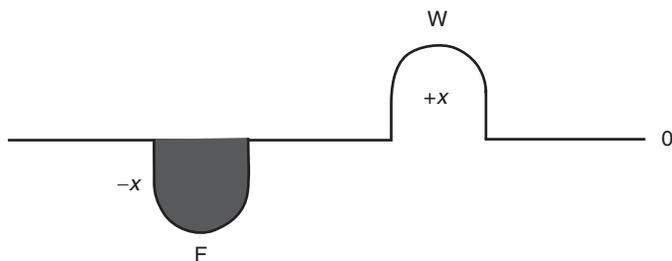


Figure 1.1 The result of the payment of wages

of wages, money is created and is given a real content. By the same token, it is transformed into income and deposited with the bank. The formation of income marks the passage from a flow (money) to a stock. Money disappears, and gives way to monetized output. Indeed, this is precisely what income is, namely the result of a transaction by which physical goods and services are injected into a numerical form, and are thus momentarily replaced by a sum of money income. Figure 1.1 illustrates what happens in this regard.

The debt incurred by F to B is represented as a negative bank deposit, whereas the positive bank deposit represents the credit of W (B's debt to W). Since the object of F's debt is nothing but produced output, we have lodged it in F's negative bank deposit. When wages are paid, the positive and negative bank deposits are no longer defined with respect to the same agent. F becomes a debtor and W a creditor to the bank. Before the payment, debit and credit have no real content, and the object of B's acknowledgement of debt is the acknowledgement of debt itself. After the payment, the object of both W's net credit and F's debit is produced output. Physically stored in the firm's warehouse, produced output is economically lodged into F's negative deposit with the bank, where it momentarily disappears. From an economic point of view, physical output is replaced by income. This is so because income – the object of W's net credit to the bank – is the very definition of produced output.

In his book on inflation, unemployment, and capital malformation, Schmitt (1984a) introduced the concept of *absolute exchange*, that is, of a transaction in which a given object is exchanged against itself, albeit in a different form. The payment of wages is actually an absolute exchange, where output is transformed into a sum of money income.

The payment of wages is an emission: this is to say that [...] workers get *their own product*, in money. This is not just an equivalence, but

an identity: each worker obtains a money that, given its emission in the payment of wages, is made identical to the real product of this very worker. In the same transaction, the firm gives and gets the same object; which is the mark of an absolute exchange.

(Schmitt 1984a: 347, our translation)

In the absolute exchange defined by the payment of wages, physical output becomes the object of *W*'s income, but it is also, at the same time, the object of *F*'s debt to the bank. Even though output is physically stored in *F*'s warehouse, it is not owned by *F*, precisely because *F* has incurred a debt to *B*, whose object is physical output. On its turn, *B* cannot be the economic owner of physical output, because *B* is indebted to *W*. Indeed, workers are the true owners of newly produced output, since they own a net credit over a positive amount of income deposited with *B*. By transforming output into a sum of income, the absolute exchange that takes place when wages are paid out gives workers the economic ownership of the product resulting from their own activity. This conclusion should not come as a surprise, because it is the direct consequence of labour's particular status. It is because labour is the only (macroeconomic) factor of production that its remuneration is at the origin of income, and it is because wages are the outcome of production that they define produced output. Through the payment of labour, physical goods and services are made identical to money wages, so that income holders are the economic owners of produced output.

Money's value, to wit, its purchasing power, is defined by those goods and services with which money is identified. The idea that money's value is what allows money to be exchanged against real goods and services is misconceived, if this is considered as a relative exchange in which money's value is the counterpart of produced output. In no circumstances it is possible to separate income from output as if they were two distinct assets. Income and output are the two faces of the same reality. They define each other, neither of them having an autonomous existence with respect to the other. Indeed, without a numerical form, output would be only a heap of physically heterogeneous objects, and without a physical content money would amount to an empty numerical form of no positive value. It is through their strict association that they acquire a meaningful standing, namely that physically heterogeneous objects are transformed into commodities, and money is given a real content. In the absence of an absolute exchange transforming physical output into money income, money's value could never be explained. The very idea of a positive purchasing power implies the

existence of a strict relationship between money and output. The payment of wages allows for money and output to become the two terms of an identity, which is the strictest possible relationship. Money's purchasing power results from such a relationship: it is the direct consequence of money income being the very (economic) definition of produced output.

Nothing of what we argued so far about the nature of bank money and the formation of income has been assumed as dogmatically given, or derived from a set of axioms. The double-entry bookkeeping principle has been the only instrument used in our analysis of the way money enters the realm of economics, and its creation is made to overlap with production. Whatever theory economists adhere to, they cannot confute the principle of double-entry bookkeeping. Hence, if economists are still far from an agreed understanding of what money is, it is because this principle has not yet been assimilated within their theoretical framework. In particular, negative numbers are still alien to their monetary writings. However, negative numbers are essential for the very existence of double-entry bookkeeping, which is based on the constant matching of negative and positive numbers. The necessary equality between debits and credits, as a matter of fact, is an identity between negative and positive numbers. Once this is clearly understood, it appears that each transaction has to be entered both on the assets and on the liabilities sides of a bank's ledger, so that the result for the bank (considered here as a simple intermediary, and not as a firm) can always be equal to zero. As odd as this might first appear, this implies that each agent intervening in a transaction is simultaneously credited and debited (or debited and credited) for the same amount of money units.

Let us refer to the double-entry representation of the payment of wages (Table 1.2). When B pays x to W on behalf of F, it credits and debits both F and W with x units of money. In the same instant, B creates and destroys x units of money on F as well as on W. Indeed, F cannot be the beneficiary of a positive creation of x money units, as they are immediately used in the payment of W. On the other hand, workers deposit at once the money they are credited with by the bank. As a result of the payment of wages, workers are net creditors not because they hold a positive amount of money units, but because they deposit their money wages with B (thus lending their income to it). If we consider money alone, we notice that, according to its nature of a numerical 'vehicle', it instantaneously flows from B to F, from F to W, and from W back to B. The accounting representation of money's circular flow requires F and W to be simultaneously credited and debited with x money units, which

implies that the money created in the payment of wages is immediately destroyed once the payment has been carried out.

The flow nature of money has for a long time remained hidden, and has yet to be perceived as a central feature of monetary analysis by most economists. Substantially, money and income have erroneously been confused for one another, and money as such has not been given a proper definition. However, it is only fair to recognize that Smith (1776/1991) had already distinguished money (nominal money) from money's worth (real money) – a distinction that was to be taken on by Ricardo (1817/1951) and by Marx (1867/1976). According to Smith (1776/1991), nominal money corresponds to what we would call today 'the numerical form', a unit of account of no intrinsic value, while real money – money income or simply income in today's language – is defined by the amount of real output money comes to be identified with once associated to real production. In his famous book on the nature and causes of the wealth of nations, Smith (1776/1991: 254) defines money as 'the great wheel of circulation', and clearly states that, as such, it must be distinguished from a country's revenue. '[S]o money, by means of which the whole revenue of the society is regularly distributed among all its different members, makes itself no part of that revenue. The great wheel of circulation is altogether different from the goods which are circulated by means of it' (ibid.: 254). Smith then clarifies that (nominal or vehicular) money shall not be identified with money's worth (real money or income). '[T]he wealth or revenue [...] is equal only to one of the two values which are thus intimated somewhat ambiguously by the same word, and to the latter more properly than to the former, to the money's worth more properly than to the money' (ibid.: 255). The Scottish author concludes by pointing out that '[t]hough we frequently, therefore, express a person's revenue by the metal pieces which are annually paid to him [...] [w]e still consider his revenue as consisting in this power of purchasing or consuming, and not in the pieces which convey it' (ibid.: 255). Even if it is true that Smith's (1776/1991) analysis of money is still influenced by an old-fashioned, metallist representation of money, it is difficult to deny that by distinguishing money from money's worth, and by conceiving of money as a circular flow ('great wheel of circulation') he provided the elements for an entirely new conception of money. Unfortunately, with the partial exception of Ricardo and Marx, his message went unheeded, and the reason for which money must be conceptually separated from income seems to be alien to today's analyses. To be sure, a distinction is made, but it is confined to the obvious consideration that money is used to

express and represent income. What is missing is the awareness that money is a mere flow or a numerical form of no value at all.

As intuited by Smith (1776/1991), money's purchasing power does not derive from any hypothetical, intrinsic value of money, but rather from its identification with produced output, which results from the payment of wages. '[T]hough the wages of the workman are commonly paid to him in money, his real revenue, like that of all other men, consists, not in the money, but in the money's worth; not in the metal pieces, but in what can be got for them' (ibid.: 260). The author of *The Wealth of Nations* was himself the victim of the generally accepted view that money had a value of its own (since it was made of precious metals); however, he was not fooled into identifying it with money's purchasing power. Today's analysis of book-entry money shows that Smith was right. Whatever material is used to represent it (from gold to electric impulse), money is not a revenue or an income. It becomes (is transformed into) an income once it is made identical to produced output. It is at the very moment wages are paid that money is created, and it is in this same instant that it acquires its purchasing power. To be precise, money as a flow exists only within a payment. Once wages have been paid, money is destroyed and is immediately replaced by a sum of income. What has to be clearly understood is that banks cannot create a positive amount of income, whose origin has to be looked for in the process of production. Things being what they are, our next step is to analyse the relationship existing between money and credit.

Money and credit

One of the most enduring and serious mistakes in economics literature is to identify bank money with credit and to maintain that banks can create a positive purchasing power. As claimed by Bernstein (1965: 47), '[n]ew money is created in response to credit expansion by commercial banks', because '[w]ith the punch of a bookkeeping machine, the total amount of purchasing power in the economy [is] obviously increased' (ibid.: 44). Let us take Parguez and Seccareccia (2000) as a more recent example of this view. In their analysis of the nature of money in the monetary circuit, they maintain that

[b]anks are deemed to be so creditworthy that no holder of their debts would ever ask for reimbursement either in kind or in the debt of another agent. Banking institutions enjoy, therefore, the capacity of freely issuing debt without it being subject to an exogenous resource

constraint. This means that banks can create these debts *ex nihilo* when they grant credit to non-bank agents who must spend them to acquire real resources.

(Parguez and Seccareccia 2000: 103)

This quotation ascribes to banks the faculty of creating a money already endowed with a positive purchasing power, which the beneficiaries of the creation can (indeed must, according to the authors) spend for the purchase of real resources. Is it really necessary to dispute such a fanciful claim? Most probably not, since it is immediately clear that banks' alleged faculty to create a positive purchasing power pertains to the world of imagination. In the real world, a creation is possible only if it is immediately matched by an equivalent destruction, a result that can be achieved once negative numbers are associated with positive numbers. This is what happens when bank money is issued in compliance with the rules of double-entry bookkeeping. To endorse the idea that banks can create credit is to evoke a supernatural power that goes far beyond our terrestrial competence.

Nevertheless, it seems difficult to give up the widespread idea that money creation is closely related to credit. Indeed, what cannot be accepted is the axiomatic assumption that banks can create credit, and not the existence of a necessary link between money creation and credit. In order to understand the origin and nature of this link, we need to take into account the role of intermediation played by banks. In particular, we have to investigate the financial intermediation carried out by banks in the payment of wages, which is – as we already know – the transaction leading to the formation of income.

Wages would be paid out of a positive income only if banks were able to create it, in which case we would be entitled to identify money creation with a credit operation. Banks would lend the newly created income to firms, which would use it in the purchase of produced output from workers. This not being the case, if a credit is nevertheless involved in the payment of wages, it must originate in the income earned by workers. As a matter of fact, as soon as wages are paid out an income is formed, which immediately takes the form of a bank deposit. As shown in Table 1.2, the sum paid to workers is entered on the liabilities side of B's ledger, which means that B has just incurred a debt to W. The reason for B's indebtedness is that it benefits from a loan: W's income, formed as a bank deposit, is lent to B at the very moment of its formation. Workers end up with a claim on a bank deposit, which states that they are the owners of the income deposited with (and thereby lent to)

B. We are thus confronted with the first part of a credit operation that has its origin in the income earned by workers and borrowed by the bank where it is deposited.

Table 1.2 helps us to understand also the second part of B's financial intermediation. Being entered on the assets side of B's balance sheet, firm F acknowledges its debt to the bank. At the same time, B balances its debt to W with an equivalent credit, whose meaning is univocal: the income lent by W to B is immediately lent by B to F. The bank's financial intermediation is thereby instantaneously completed, and the credit accompanying the payment of wages acquires its full meaning. F benefits, indeed, from a credit provided by B, yet the income lent to F has not its origin in B, but in W: it is the wage earners' income that feeds the bank's credit to firm F. The role of B is to transfer to F the income earned by W. Thanks to B's financial intermediation, the firm can now purchase on the factor market the goods produced by its workers. It is important to note, however, that this is not a final purchase. F will have to reimburse B, and it will be through the final sale on the goods market of its workers' output that it will be able to meet its obligation. Since F's purchase is financed by W's income, the only apposite conclusion is that the wage earners are still the owners of produced output. At the same time, it is also true that F's purchase is an investment, and that, in capital terms, produced output is a stock owned by the firm. Hence, while F is entitled to handle and sell current output, it does not truly own it, since output is the very object of F's debt to the bank.

In the payment of wages, the bank intervenes both as a monetary and a financial intermediary. On the one hand, it issues the numerical vehicle necessary to convey the payment and to provide the numerical form for physically heterogeneous goods and services to be made homogeneous. On the other hand, it transfers to F the income newly formed by workers. Monetary and financial intermediations take place simultaneously, so that it is finally acceptable to claim that by paying wages to F's workers B grants a credit to F, provided that it is clearly stated that the income F is credited with by B is W's income, that is, the very income formed in the payment of wages. It thus appears that, through B's financial intermediation, F benefits from a credit granted by W. The creation of money by a bank does not imply any creation of income, and cannot be assimilated to an ordinary credit operation.

Let us consider in this light the analysis of bank money advocated in the monetary circuit approach. As shown by the following quotation, the authors supporting this approach claim that, through money creation, banks credit firms with a positive income created *ex nihilo* that

enables them to purchase whatever item they need, and that the credit granted by banks through money creation is substantially of the same kind as the credit financed out of accumulated savings.

First, money is, and has always been, a debt created *ex nihilo* by bank credit advances that are granted either to permit the generation of real wealth or to acquire existing physical assets. Second, there is no alternative between debt financing on the one hand, and the tapping of existing liquid resources (or accumulated savings) to 'finance' expenditures, on the other.

(Parguez and Seccareccia 2000: 102)

What Parguez and Seccareccia (2000) do not seem to be aware of is that, within the payment of wages, the credit accompanying money creation is substantially different from ordinary credit (financed by accumulated savings). When a bank lends a sum of income previously formed, saved, and deposited with it, it simply transfers an existing income from savers to borrowers. A money is issued, which conveys this transaction, yet it is an emission that does not end up in the formation of any new income. An ordinary credit is a microeconomic transaction that leaves the amount of income available in the whole system unaltered: 'the transaction on the financial market is an ordinary credit for the simple reason that the lender loses a positive money [income] to the benefit of the borrower' (Schmitt 1984a: 156, our translation). Things are different when credit relates to the transaction defining the formation of a positive new income. In the payment of wages, income is not the prerequisite, but the result of the transaction. The income lent to firms is the very same that is formed when wages are actually paid out and this credit is not an ordinary but a quantum one.

The emission of wages is not an operation of credit; it is therefore incorrect to assimilate bank money to a credit or to a credit money. Nevertheless, if we still want to speak of credit when referring to the emission of wages, we must specify that we are confronted with an entirely original credit, a *quantum credit*, creation and not transmission of money [income].

(ibid.: 154, our translation)

In this brief analysis on the nature of bank money, a last point about the relationship between money and credit must be clarified, namely the alleged capacity of banks to multiply their deposits. The money

multiplier is accepted without reserve, taught in most bachelor programmes of economics, and expounded in every textbook on monetary theory. The idea is very simple. Let us imagine that an economic agent, A, deposits her/his newly formed income, equal to 100 money units, with a bank B, and suppose that B has to deposit 20 of these 100 units with the central bank as compulsory reserves. According to the money multiplier thesis, bank B will lend 80 units to another economic agent, C, who will spend them in favour of D, thus increasing D's deposits with her/his bank (which, for the sake of simplicity, we identify with B). If the central bank's policy remains unaltered, B will then grant another credit of 64 money units to client E, who will use it to finance her/his purchases from F, whose bank deposit will increase by the same amount. The process will go on like this until the sum of B's compulsory deposits with the central bank reaches the amount of A's initial deposit (100 money units). On the whole, B will end up having granted credits for a total of 500 money units (inclusive of the initial 100 units). We are thus led to conclude that, owing to this simple mechanism, the initial deposit may be multiplied by 5. Theoretically, if no compulsory deposit with the central bank was required, the process would never end and money would be multiplied *ad infinitum*.

Next of kin to Kahn's income multiplier, the money multiplier is a clear example of a monetary illusion owing to a poor understanding of the nature of bank money and income. A rigorous analysis of our simplified example should be enough to show that no miraculous multiplication of money or deposits occurs in the real world. As far as the first deposit is concerned, it should be clear that it defines the income earned and saved by A. This means that a production must have occurred, which explains the formation of 100 units of money income. The subsequent loan of 80 money units to C defines the amount of A's income transferred to C. Now, the sum borrowed by C can be spent for the purchase of either the output whose production had initially led to the formation of income, or of the output of some other production. In the first case, the income spent by C is simply destroyed. In the second case, the new sum of 80 units deposited by D is not due to a phantasmal multiplication by 0.8 of the initial deposit, but to D's own productive activity. If the first production takes place in period 1, and the second in period 2, for the two periods taken together bank deposits amount to 180 money units. There is nothing surprising or mysterious in this result. Every production creates an income that defines it, and which is deposited with the banking system. It is not the bank that, by lending its deposits, increases their amount.

When C spends A's income for the purchase of D's output, s/he does not create a new income in the hands of D. Two scenarios are possible. In the first, D is a firm, and its output is the making of its workers whose wages amount to 80 money units and are deposited with B. Before C's expenditure of the sum borrowed from B, total bank deposits amount therefore to 180 money units. The final purchase by C of period 1 output reduces the amount of income to 80 units, the very sum defining period 2 output, and required for its final purchase. In the second scenario, D is a self-employed worker who sells its own product to C. Even though there is no explicit payment of wages referred to D's production, it would be mistaken to think that no new income defining D's output is formed. When C purchases D's product, her/his payment has the double effect of destroying an income and forming another one. The income spent by C is destroyed, since it finances the final purchase of an equivalent output, and is replaced by a new income generated by the production carried out by D. Agent D is simultaneously a firm and its worker, and C's payment defines at the same time a final purchase and a payment of wages. At the end of period 2, the income still available is of 100 money units – 80 units deposited with B and 20 units deposited with the central bank – whereas the totality of B's deposits for the two periods considered together is equal to 180 money units. Finally, both scenarios lead to the same result: the increase in B's deposits is due to production, not to a process of multiplication that banks would engender simply by lending their deposits.

To avoid endorsing such an extraordinary claim as that implied by the money multiplier theory, we must revert to the principles of double-entry bookkeeping, and to the necessity to balance each credit with an equivalent debit of the same agent. Further, it should not be forgotten that banks intervene as simple intermediaries. This means that a credit may be granted only provided banks benefit from a positive deposit. If such a deposit does not pre-exist or is not created in the payment carried out by banks on behalf of their clients, it is vain to look for a transaction allowing banks to create a positive purchasing power out of the blue.

Every time money that has been deposited is re-lent [...] this process is to be regarded as the creation of additional purchasing power; and it is merely this comparatively simple operation that is at the root of the banks' ability to create purchasing power – although the process appears so mysterious to many people.

(Hayek 1933/2008: 26)

This hypothetical process does not only appear to be mysterious, it is right away impossible actually.

For a long time, economists have hopelessly tried to reconcile their belief in the faculty of banks to create credit with the claim of bankers that loans can be granted only if they are financed by equivalent deposits. If banks could finance final purchases through money creation, they would also increase their deposits, and thus respect the rule of the necessary equality between loans and deposits. This is the argument that apparently sets the question. However, this is to forget both that income is formed by production, and that final purchases imply its destruction and, therefore, the destruction also of the bank deposits defining it.

Let us consider once again what happens in Table 1.2. The formation of a positive income deposited with B is not the result of any particular theory chosen according to a set of more or less acceptable hypotheses. Double-entry bookkeeping is not a theory, which can be supported or invalidated by facts. The payment of wages takes the form represented in Table 1.2 irrespective of any theoretical choice or drive of the researcher. Analogously, the final expenditure of income in the purchase of current output can only be represented as shown by Table 1.3.

Although wage earners are the initial holders of newly formed income, they are bound to transfer part of it to other economic agents – such as the State or the firm (when profits are formed) – which is why we have replaced workers with income holders. If we regroup Tables 1.2 and 1.3 into a single table, and we account for the distribution of income from workers to income holders (of which workers are still part), we obtain Table 1.4.

We immediately observe that a compensation occurs between F's debit and credit, as well as between income holders' credit and debit. Whatever our theoretical standpoint may be, double-entry bookkeeping forces us to conclude that the final expenditure of income leads to its destruction, that is, to the disappearance of the deposit initially formed with B.

Table 1.3 The purchase of produced output

Bank B			
Assets		Liabilities	
Credit on wage earners	x m.u.	Debit to firm F	x m.u.

Table 1.4 The result of the final purchase of output

Bank B			
Assets	Liabilities		
Credit on firm F	x m.u.	Income holders' deposit	x m.u.
Credit on income holders	x m.u.	Debit to firm F	x m.u.
	0 m.u.		0 m.u.

If we were now to endorse the claim that banks create credit to finance final purchases, we would have to add a sum of money (y) to that of the income formed in the production of current output. As before, the final purchase of output would define the destruction of x units of money income. The y units of money added by the bank would remain deposited with it, to the benefit of F, and the consumers, beneficiaries of B's credit, would balance B's debit to F. However, neither F's credit nor the consumers' debit would have a real object. Current output, the sole product in our example, has been sold, which leaves the y units of money deposited with B with no real content whatsoever. Far from defining a creation of purchasing power, the credit granted by B in order to finance the purchase of current output would amount to an emission of 'empty' money, and would define an inflationary increase of money units. What is presented as the normal working of a banking system, in reality is an entirely anomalous process that has nothing to do with the true nature of money and banks. To be true, anomalies are also part of today's monetary system, but in order to understand their origin it is first necessary to expound the logical laws that govern the system. To maintain that anomalies are the rule is to turn the problem upside down, trying to avoid strict logic. As we shall see in the following chapters, logic must guide us in an analysis that has first to be positive, before being normative. It is positive analysis that, by determining the rules inherent in the nature of any economic system, provides the litmus paper required to detect the presence of an anomaly and to discover its pathogenesis.

2

The Macroeconomic Laws of Monetary Production Economies

In this chapter, we will show that, when referred to money and income, economic laws are no less rigorous than physical laws. Their validity is not influenced by economic agents' behaviour, since they belong to the category of identities, and not to that of equilibrium conditions. Whatever decision may be taken by individuals or institutions, these laws remain unaltered. They set the logical rules at the roots of any economic system based on bank money. As such, they do not refer to any specific economic policy, but rather to the monetary structure on which our economic systems are founded. In other words, they pertain to the logical infrastructure making up for the conceptual and practical framework within which economic agents are free to take their decisions. Economists have argued for a long time about the primacy of identities over conditions of equilibrium, and vice versa. Following Walras's (1874/1984) contribution, and despite Keynes's (1930/1971) suggestions, it is today almost generally accepted that economic laws are mainly behavioural, and founded on microeconomics. Mathematics is widely used, and equations of various kinds and complexity are considered as one of the best tools available to represent the real world of economics. Identities are rarely referred to, and when they are it is to transform them almost immediately into conditions of equilibrium. Our aim is to re-establish the logical priority of identities and show that they are the very foundation of macroeconomic analysis. We will do so by starting from a reappraisal of Say's law, and a reinterpretation of Keynes's identity between global demand and global supply, to end with Schmitt's law of the necessary equality between each agent's sales and purchases.

Say's law: A reappraisal

Say's main economic work, the *Traité d'économie politique* (1803/1972), has been the object of numerous interpretations. Leaving aside any attempt to establish what Say's own thought really was, we will emphasize only a few aspects of his contribution, with the aim to provide for an enunciation of his law that proves useful in our effort to unveil the nature of money and income. The expression 'supply creates its own demand' is not to be found in Say's book. It was proposed by Keynes (1936/1946) in his *General Theory*, and, even though it has not been unanimously accepted, is generally referred to as Say's law. Now, what really interests us here is not whether Say would have himself used this expression or not, but what is the meaning it conveys, and how this meaning is supported by Say's analysis.

As often is the case when a contribution is considered important and innovative, Say's work has been interpreted in the light of mainstream economics with the intent to show that the former supports the latter. It so happens that Say's contribution has been praised by neoclassical economists, who see in the French economist's claim that 'products are paid for with products' (Say 1803/1972: 125, our translation) the clear indication that he advocated the idea of relative exchange. Hence, according to Hutt's 1974 book on the rehabilitation of Say's law, the claim that goods and services exchange for goods and services 'is a very broad yet apt statement of the "laws of markets" [Say's law]' (Hutt 1974: 6). However, to say that products are paid for with products is not tantamount to claiming that products exchange directly for products. As shown by the following quotation, the idea conveyed by Say's sentence is that money is an intermediary, and that, once it has fulfilled its role, it disappears, so that, finally, payments are all real. 'Money performs a momentary function in this double exchange; and when exchange comes to an end it will always be found that products have been paid for with products' (Say 1803/1972: 125, our translation).

The crucial point here is to determine whether – as maintained by general equilibrium analysis – money is a simple veil, which does not alter the terms of relative exchange, or an intermediary that, through its association with produced output, sets the term of exchange. In both cases, products are eventually paid with products. In the first case, money plays no determinant role, whereas in the second case it is all-important. If money is a veil, direct (relative) exchange is what really matters. Everything would happen as if real goods and services were directly exchanged for one another, in a series of transactions that would

determine their relative prices. If money is an intermediary acting as a numerical form and as a flow conveying payments, products are first the object of a series of absolute exchanges determining their numerical prices. Through production, money is immediately transformed into income, and the expenditure of income conveys real goods and services to income holders.

The choice between the two alternatives reported above is not arbitrary. Money is not what the advocates of a given theory want it to be. Its nature is univocally defined, and today's monetary systems allow for its clear identification in the form of bank money. Say's investigation was more arduous, since the use of gold certainly did not help the understanding of money's immateriality. It is thus all the more remarkable to find out that he was able to avoid both the mistake of identifying money with a commodity, and that of considering it a mere veil. The following quotation puts it very clearly.

A priest goes to a merchant to buy a stole or a surplice. The value he brings with him is in the form of a sum of money. Who gave it to him? The tax-collector who in turn had levied it from a tax-payer. Who gave it to the tax-payer? He had produced it himself.

(*ibid.*: 124–5, our translation)

Obviously enough, what Say is referring to in this quotation is money income, and he is uncompromising in asserting that (1) it is money income that finances final purchases, and that (2) income is formed through production. Although it remains true that, eventually, products are paid with products, payments are conveyed by money and financed by income. Production precedes exchange on the commodity market, and money is introduced since the process of production. It is here that it is associated with produced output, and, by means of an absolute exchange, transformed into income.

This interpretation of Say's analysis is corroborated by the definitions he gives of production and consumption, which are respectively identified with the creation and the destruction of value (Say 1803/1972: 128). If we keep in mind that value is used by classical economists to indicate money's worth, it is perfectly licit to infer that Say considered income (money's worth) to be the result of production and identified consumption with the final expenditure of this same income. The applicance of the concepts of creation and destruction to production and consumption is the mark of Say's originality and shows how innovative his analysis was. In the light of Say's conception of production and consumption, his law may be formulated by saying that production

creates the exact amount of income required for the final purchase of current output or, in the words of Sowell (1994: 39), that '[t]he total factor payments received for producing a given volume (or value) of output are necessarily sufficient to purchase that volume (or value) of output'.

As is well known, Keynes rejected Say's law on the ground that it does not allow to explain general unemployment. In Chapter 3 of his *General Theory*, Keynes (1936/1946: 26) writes that 'Say's law, that the aggregate demand price of output as a whole is equal to its aggregate supply price for all volumes of output, is equivalent to the proposition that there is no obstacle to full employment'. It is our contention here that Keynes's criticism does not apply, since it refers to a pathological state of the economy. Say's law does not give any hint about the cause of unemployment, it is certain, but this only means that Say's analysis has to be completed by a theory showing that the necessary equality between global supply and global demand may be consistent with deflation (and inflation). That Say's analysis is not up to this task is a fact. However, this does not imply that his law is mistaken or useless. Quite to the contrary, Say's law is a logical identity, which sets the rule any correct theory must comply with. It is a necessary cornerstone, providing the fixed point required to compare a disorderly with an orderly system. In its absence, pathology itself could not be defined, let alone explained. From this point of view, Keynes's own analysis is much less at odds with Say's than most economists are led to think, and we can share Kaldor's claim that Keynes's theory 'is best analysed as a development or refinement of Say's Law, rather than a complete rejection of the ideas behind the law' (Kaldor 1983: 6).

Keynes's identity between Y and $C + I$

At the beginning of Chapter 10 on the fundamental equations for the value of money of his *Treatise on Money*, Keynes introduces the crucial concepts on which his identities are founded. In particular, he starts by individuating in production (human effort) and consumption the determinant factors of every economic system. 'Human effort and human consumption are the ultimate matters from which alone economic transactions are capable of deriving any significance; and all other forms of expenditure only acquire importance from their having some relationship, sooner or later, to the effort of producers or to the expenditure of consumers' (Keynes 1930/1971: 120–1). He then goes on by specifying that both production and consumption concur in the definition of income, of which they represent the two joint aspects.

I propose [...] to start [...] with the flow of the community's earnings or money income, and with its twofold division (1) into the parts which have been *earned* by the production of consumption goods and of investment goods respectively, and (2) into the parts which are *expended* on consumption goods and on savings respectively.

(*ibid.*: 121)

Keynes's first identity follows from his identification of production and consumption as the two terms of an identity, and takes the well-known form of

$$Y \equiv C + I,$$

where Y stands for national income, C for global or total consumption, and I for global or total investment. 'Income = value of output = consumption + investment' (Keynes 1936/1946: 63). Income results from production and is spent for the final purchase of consumption and investment goods. The same object is apprehended from two different points of view. It would be mistaken to claim that Y and $C + I$ are two distinct elements, which can differ from one another. In reality, income is defined by the unity of Y and $C + I$. It thus appears that Keynes's identity between Y and $C + I$ establishes the necessary equality between global supply and global demand, which is nothing but another version of Say's law. Current production creates an output whose value is the income necessary and sufficient for its final purchase in the form of consumption and investment goods: supply creates its own demand.

Economic theories are traditionally distinguished according to whether they privilege supply or demand. Thus, neoclassical theories are ranged in the field of supply-side economics, while Keynesian theories pertain to the field of demand-side economics. This fosters the idea that it is either supply that determines demand, or demand that determines supply. In both cases, supply and demand are seen as two distinct entities, whose equivalence is verified only for a specific equilibrium level of either prices or income. What Say's law and Keynes's identity tell us instead is that supply and demand are jointly determined. Current output and income are not two distinct results of production. If they were, total wealth would be defined as the sum of income and produced output, a hypothesis already rejected by Smith (1776/1991: 255), who correctly observes that '[t]he whole revenue of all of them taken together [all the inhabitants of any country] is evidently not equal to both the money and the consumable goods; but only to one or other of those two

values'. Production cannot give rise simultaneously to current output and to a distinct and equivalent amount of income, unless we are ready to admit that monetization has the double effect of providing a numerical form to physical output, as well as of adding a new value to it. The risk of counting the same value twice is avoided once it is understood that production has a unique result, whose two joint aspects are the terms of an identity. Current output, the supply side, is identical to current income, the demand side, and it is vain to look for any quantitative difference between them.

The determinant point is that production is not a relative but an absolute exchange, that is, an exchange that gives another form to current output. Through the payment of wages, physical output is transformed into a sum of money income. This is what determines Y , the first term of Keynes's identity. The second term relates to demand. It might be thought that demand must be referred to final consumption, so that a positive period of time separates supply from demand, which cannot be considered as the terms of an identity. This is to forget that demand is financed by income. Without a positive income, demand can only be nil. Moreover, income is formed as a bank deposit, and bank deposits resulting from production are immediately lent to firms. The income formed through the payment of wages is invested by firms in the purchase of current output, which is thus transformed into a stock. F 's investment is a true demand, exerted simultaneously with the formation of income.

Keynes could not provide the full proof required to corroborate his intuition. The concept of absolute exchange was alien to him, as was Schmitt's (1984a) quantum analysis. Not surprisingly, Keynes's followers have considered his identity between Y and $C + I$ as a relationship between two distinct terms that may be equal at equilibrium only. Chronologically, the instant when income (Y) is formed precedes the instant when income is spent for the final purchase of consumption and investment goods ($C + I$). However, the identity between Y and $C + I$ requires their simultaneity. How can these two apparently contradictory requirements be reconciled? The answer rests on Schmitt's (1984a) discovery of quantum time. Although separated in chronological time, production and consumption coincide in quantum time, since they are instantaneous events defining the same quantum of time. The fact is that consumption concerns the final purchase of the same output whose production is at the origin of income. Since production and consumption define the creation and the destruction of the same income, and since income is the very definition of output, at the instants of

production and consumption we observe the complementary transformation of the same output. When income is formed, physical output takes the form of money; when income is destroyed, output recovers its physical form. Taken together, creation and destruction are the two complementary faces of a unique emission: *'every production and (identically) every expenditure is an emission, creation-destruction of the same object, the product and the expended sum'* (Schmitt 1984a: 76, our translation). The concept of quantum time relates to this emission, positive and negative, of the same output, and to two well-established results of monetary macroeconomics. The first is that every output is formed instantaneously but relates to a finite period of time (its period of physical production). *'Production quantizes time; that is to say, it takes hold in an instant of a portion of continuous time: the first result of production is therefore the definition of a quantum of time. Output is not set in time; it is time'* (ibid.: 54, our translation). The second is that consumption is but the final purchase of produced output, that is, an instantaneous event (expenditure) relating to a finite span of time. *'Expenditure is a unique event, even though it relates to a whole period of time; it is an instantaneous magnitude related to a positive period'* (ibid.: 66, our translation). Production amounts therefore to the emission (the instantaneous definition) of the same period of time that is emitted when consumption takes place. The quantum time dimension of production and consumption is exactly the same, which is the reason why it is correct to claim that these two events coincide in quantum time.

Schmitt's (1984a) quantum analysis is far too articulated to be resumed in a few lines, as well as far too difficult to be immediately grasped by those still unfamiliar with his ideas. The readers who are willing to go to the core of the problem are encouraged to do so by referring to Schmitt (1966, 1972, 1975, 1984a, 1996a). As far as our analysis is concerned, the considerations about the flow nature of bank money, the process by which money is associated to physical output, and the intermediations carried out by banks should be enough to establish the validity of Say's law and Keynes's identity between global supply and global demand. The study of another macroeconomic law discovered by Schmitt (1975) provides further support to Say's and Keynes's contributions.

The identity between each agent's sales and purchases

Sales and purchases are transactions involving both money and income. Their analysis is thus strictly influenced by the nature of these two

elements, as well as by their logical relationship. For example, if money were to be identified with a commodity, sales and purchases would pertain to the category of relative exchanges, and would define two distinct transactions. The immaterial nature of bank money prevents us from considering money as a commodity, and sales and purchases as relative exchanges. What has still to be established is that, because of the flow nature of money, every single economic agent cannot purchase without selling or sell without purchasing for the same amount and simultaneously.

To claim that an agent's purchases are necessarily another agent's sales is a truism or a platitude. On the contrary, the assertion that each agent's sales are always immediately matched by equivalent purchases of this same agent has to be established on logical grounds. In order to do so, let us first observe that, though conveyed by money, purchases are financed by income. A simple vehicular money has no purchasing power at all, so that it cannot finance any purchase unless it is transformed into income. The apposite question now, is how can our economic agent get hold of a positive income. The answer is straightforward: either by working, by getting indebted, or by selling a real good. In each of these cases, she/he becomes an income holder through a sale: of a service, a financial claim or a commodity. The fact that our agent can finance her/his purchases only through an equivalent sale is however not enough to establish Schmitt's law. What is still missing is the simultaneity of these two transactions. To state that in order for an individual to purchase he needs to sell is not tantamount to claim that every time that he purchases he also sells, and that his sales and his purchases are exactly of the same amount. The reader may well be prepared to accept that purchases are financed by sales, but she/he is not necessarily keen to accept the idea that these two transactions take place at the same instant. Yet, as the following analysis shows, this is the only possibility consistent with the flow nature of bank money. The fact is that money is not an asset and cannot be the final term of any exchange. In the words of Schmitt:

Since money is an asset–liability, every purchase is necessarily financed by a sale, money being only an intermediary in the transaction. Likewise, a sale cannot have money as its final object, since the product that is sold is a net asset, whereas money is no asset at all. For the transaction to be completed it is therefore necessary to compensate the purchase with a sale. This law can also be formulated by saying that every monetary transaction is half a real

transaction, either a purchase or a sale, completed by the reciprocal half transaction.

(Schmitt 1975: 19, our translation)

Let us recall once again the example of the payment of wages considered in Chapter 1. The payment of wages amounts to the sale, by workers, of their labour services. If this sale is compared with the final purchase of produced output by the same workers, Schmitt's law would not be corroborated, since the sale of labour services is not simultaneous with output's final purchase, and since workers are likely to spend less than what they have earned (their income being reduced by taxes and other transfers). It is clear, therefore, that the identity between workers' sales and purchases cannot be verified on the labour and on the commodity market alone. But these are not the only markets available. Another important market has to be taken into account, whose role is important in our economic systems: the financial market. At this point, we have to examine what happens on the financial market at the very moment wages are paid to workers. Once again Table 1.2 helps us to apprehend the implication of double-entry bookkeeping. The fact that workers are entered on the bank's liabilities side signifies that they own a credit, or, in other words, that they hold a financial claim whose object is the income deposited with the bank. The immediate deposit of wages with the bank makes the workers the holders of a deposit certificate, which is best defined as the purchase of a claim on the financial market. As shown by double-entry bookkeeping, workers' sale of labour services is thus balanced at once by their equivalent purchase of deposit certificates.

The relationship between workers' sales and purchases is a true identity, since it is impossible for wages not to be deposited with the bank. The payment of wages creates a bank deposit to the benefit of workers. Income earners are credited with a deposit certificate, which is the best evidence that they spend at once what they obtain through the sale of their labour services. It is in the nature of bank money to be a flow. Hence, money is always used in an instantaneous circular flow, which has the bank as its point of departure and arrival. Workers can never hold a positive amount of money units. They are simultaneously credited and debited with the same money units. Double-entry bookkeeping is univocal: workers are entered on the liabilities side of the bank's ledger because they cannot get hold of the money units conveying the payment of their wages. Likewise, workers cannot take hold of the income formed at the moment they are paid for their labour services.

The reason is that the payment of wages is a three-polar transaction where the bank acts both as a monetary and a financial intermediary. Workers cannot avoid depositing their income with the bank, which lends it immediately to firms. This does not mean that workers get nothing in payment of their labour services. On the contrary, following the payment of wages, workers become the owners of produced output in the form of money. However, instead of holding it in this form, they give to their economic right the form of a financial claim. They lend their income to the bank and get a financial claim in exchange.

As a further example, consider now what happens on the side of firms. They are the beneficiaries of the bank's loan, and thus become its debtors. This is confirmed by the fact that firms are entered on the assets side of the bank's balance sheet. But to be granted a loan is tantamount to becoming indebted, which is tantamount to selling a financial claim. In order to obtain a loan, I must accept to sell an acknowledgment of debt. This is what happens to firms: they are credited with the income saved by workers and they give the bank an equivalent amount of financial claims, whose object is physical output. Yet, firms spend at once the income lent to them by the bank. They do so in order to purchase current output and transform it into a stock. Firms' sale of financial claims is hence accompanied by their purchase of current output. As a result, firms own a stock of goods and are indebted to the bank for the very amount of income defining these goods. This means that, although they hold it in the form of a stock, firms are not the final owners of produced output, which they have to sell in order to cancel their debt to the bank. As far as Schmitt's law is concerned, it is verified by the simultaneity of firms' sales (of financial claims) and purchases (of produced output).

To conclude, let us corroborate Schmitt's law with a brief analysis of the transactions taking place when income holders finally purchase produced goods and services. As explained above, income is immediately deposited with the bank, and income earners become the owners of claims on bank deposits. When they spend what they previously saved and lent to the bank, income earners are debited by B, which enters them on the assets side of its ledger (see Table 1.3). This entry is 'simple' and seems in contrast with the thesis that consumers are at the same time credited and debited. In reality, this is exactly what happens, even though income holders are not entered twice in the bank's balance sheet. Since wage earners hold claims on a bank deposit, the cancellation of their bank deposit (see Table 1.4) implies both that of their claims on a deposit and that of the income they still own in this form. This is to say that, in the same instant, they sell their claims on a bank

deposit and spend the income they obtain from this sale. The purchase of produced output by income earners is financed by the sale of their financial claims, in conformity with the law of the necessary equality of each agent's sales and purchases.

Schmitt's law is the most advanced and modern version of Say's law and is the milestone of Schmitt's analysis of the circuit applied to national and international payments. We shall not enter into this analysis at this stage, but confine ourselves to observing that, consistently with Say's law and Keynes's identity, Schmitt's law concurs in providing the logical basis for a correct understanding of the way our economic systems work. The reader looking for empirical confirmation might well be surprised by this conclusion. Is it not true, she/he might claim, that our economies are characterized by the presence of recurrent crises? And is this no clear evidence that none of these laws apply in the real world? As we will show in the next section, this is not so. Contrary to the impression that would result from a superficial analysis, the identity between global demand and global supply is not only perfectly compatible with the existence of such pathological states, but proves to be necessary for the very definition of these pathologies.

Keynes's identity between saving and investment

The crucial point at stake is the concept of equilibrium between saving (S) and investment (I). Economists who assume microeconomic foundations of macroeconomic analysis emphasize the role played by factors influencing economic agents' decisions. Their search for equilibrium between S and I is therefore based on the discovery of those values of income and/or interest rate for which the consumers' decision to save is matched by the firms' decision to invest. Consumers' behaviour is determined independently of producers' behaviour, so that an adjustment is required, which, through variations in income and/or interest rate, brings saving and investment into equilibrium.

Keynes himself was not immune from the influence exerted by microeconomics. The following quotation is a well-known example of Keynes's difficulty in getting rid of old ideas: '[T]he amount of saving is an outcome of the collective behaviour of individual consumers and the amount of investment of the collective behaviour of individual entrepreneurs' (Keynes 1936/1946: 63). However, this traditional approach to saving and investment is certainly not what distinguishes Keynes's analysis from general equilibrium analysis. Even the claim that, for Keynes, saving and investment adjust through a variation of income

instead than through a variation of the interest rate does not account for the originality of Keynes's macroeconomic approach. What really marks the novelty of Keynes's message is the concept of identity, and it is by reference to this concept that the equilibrium between S and I has to be tested logically speaking.

A crucial aspect of both saving and investment is that they are (monetary) expenditures. As such, they are strictly dependent both on the nature of money and on the relationship existing between money and current output. At a first approximation, saving is that part of income that is not spent for the purchase of current output. Now, saved income defines a bank deposit owned by savers. As double-entry bookkeeping shows, bank deposits are entered on the liabilities side of banks' ledgers, and they are necessarily matched by an equivalent entry on the assets side of the same ledgers. The meaning of this double entry is clear and univocal: the income deposited with banks is instantaneously lent. In the case at stake, the income saved is lent to firms, which need it to cover the costs of production of that part of current output not yet sold to consumers.

At this stage of the analysis, it proves useful to bring in Schmitt's law of the identity between each agent's sales and purchases, a corollary of double-entry bookkeeping. Let us first apply it to the moment production takes place. Firms are entered on the assets side of banks' ledgers, because they are indebted to banks, their debt being the consequence of the payment of the costs of production (wages) carried out by banks on their behalf. In other words, firms sell their acknowledgment of debt to banks in exchange for the payment of wages to their workers. At the same time, in accordance with Schmitt's law, firms purchase the product of their wage earners by spending the income lent to them by banks and formed as a deposit when wages are paid out. Everything happens simultaneously: wages are paid to workers, a new income is formed as a bank deposit owned by wage earners, the deposited income is lent to firms, which spend it for the purchase of produced output. Now, this purchase, financed by the income saved by wage earners, is nothing else than an investment, the first investment carried out by firms.

There is a passage of Keynes's *General Theory* that is crucial for the understanding of the relationship between saving and investment. Let us reproduce it completely.

The prevalence of the idea that saving and investment, taken in their straightforward sense, can differ from one another, is to be explained, I think, by an optical illusion due to regarding an

individual depositor's relation to his bank as being a one-sided transaction, instead of seeing it as the two-sided transaction that it actually is. It is supposed that a depositor and his bank can somehow contrive between them to perform an operation by which savings can disappear into the banking system so that they are lost to investment, or, contrariwise, that the banking system can make it possible for investment to occur, to which no saving corresponds.

(Keynes 1936/1946: 81)

This quotation conveys very important ideas. Keynes starts by defining the belief in a possible difference between saving and investment as an 'optical illusion' and immediately adds that this deceptive appearance results from regarding saving as a one-sided transaction. According to Keynes, saving is instead a two-sided transaction, which leads us to think that, even if only implicitly, the author of *The General Theory* was aware of the necessity to respect the principle of 'double' double-entry bookkeeping and its corollary, namely the identity of each agent's sales and purchases.

To claim that a saver's relation to her/his bank is a two-sided transaction amounts to saying that she/he is simultaneously credited and debited, which is consistent with her/his sales being immediately matched by her/his equivalent purchases. This is precisely what happens when saving is first formed. Indeed, it is at the very moment that an economic agent earns a positive amount of income that saving makes its first appearance on the scene. When wage earners are paid for their production of current output, their banks record their wages: this deposit defines the formation of a positive saving. Hence, savers are simultaneously sellers in the labour market and purchasers in the financial market: they sell their labour services and they purchase a claim on bank deposits.

The other crucial ideas advocated by Keynes in the previous quotation are that (1) savings do not disappear into the banking system and are therefore never lost to investment and that (2) no investment can occur without a corresponding saving. Once again it is double-entry bookkeeping that accounts for the logical impossibility of saving to disappear into the banking system, as well as for its immediate transformation into investment. The necessary equality between banks' assets and liabilities makes it impossible for banks to avoid lending the totality of income deposited with them. The existence of reserve requirements is no obstacle to this result, and the latter cannot be dismissed by observing that banks may well refuse to grant credit to some of their clients. Bankers

are free to decide to whom they lend the income deposited with them, once they have the minimum reserves imposed by monetary authorities. However, whatever their decision, the entire sum deposited – and entered on the liabilities side of their balance sheet – is lent to the economic agents entered on the assets side of their ledgers necessarily.

Let us consider once again what happens at the moment of the payment of wages. As we already know, when wages are paid, an income is formed, which is deposited by wage earners and is matched by the debt incurred by firms (see Table 1.2). This means that the income saved by wage earners is lent to firms. Now, the income firms are credited with is immediately spent by them for the initial purchase of what has been produced by their workers (wage earners), that is, in order to cover the cost of production of current output. This initial expenditure of firms, which is financed through the income lent to them by banks, is nothing other than an investment, thus proving the correctness of Keynes's insight. Once it has been shown that saving finances investment, it is easy to confirm the second part of Keynes's claim, to wit, that investment requires a corresponding saving. As a matter of fact, investment is an expenditure, and, like any other expenditure, has to be financed by an income. Since income is formed by production, and since that part of current income which is spent for the final purchase of current output is what defines consumption, the income available to finance investment can only be the one that has not been consumed, that is to say, the one that has been saved.

The continuation of Keynes's quotation substantiates his analysis, in particular by reiterating the fact that to invest is to spend an income that has first to be formed and saved.

But no one can save without acquiring an asset, whether it be cash or a debt or capital-goods; and no one can acquire an asset which he did not previously possess, unless *either* an asset of equal value is newly produced *or* someone else parts with an asset of that value which he previously had. In the first alternative there is a corresponding new investment: in the second alternative someone else must be dis-saving an equal sum.

(*ibid.*: 81–2)

Of the two cases described by Keynes, the first relates to the situation we have been discussing so far, namely the investment of the income generated by a new production, while the second concerns a simple substitution between savers. The first case is by far the most interesting,

since it leads straightforward to the conclusion that no difference can exist between saving and investment. 'It follows that the aggregate saving of the first individual and of others taken together must necessarily be equal to the amount of current new investment' (ibid.: 82).

What has been established so far is that the identity between saving and investment is certainly respected at the moment production takes place, when output is given a monetary form and income is formed. What remains to be investigated is the nature of the relationship between saving and investment after consumption has taken place and firms invested in the formation of fixed capital. Is the relation between S and I still defined as an identity? And if so, how is it possible to reconcile the identity of S and I with the freedom of consumers and firms to take their decisions autonomously? Once again, Keynes points out the principle that leads us to the solution of this apparent dilemma. 'The reconciliation of the identity between saving and investment with the apparent "free-will" of the individual to save what he chooses irrespective of what he or others may be investing, essentially depends on saving being, like spending, a two-sided affair' (ibid.: 84).

Correctly understood, the principle of double-entry bookkeeping leaves no room for any difference between saving and investment in real time. This is because macroeconomic saving – the part of current income that will never be spent in consumption – is defined simultaneously by the non-expenditure of a given income in the commodity market and its investment on the labour market. Keynes has no hesitations about it: saving is an expenditure. The old-fashioned idea that saving might be compared to hoarding has nothing to do with Keynes's monetary analysis, which has bank money at its core. Double-entry bookkeeping is what really matters in this regard, and according to this principle it is logically and factually impossible for a bank deposit to disappear and for an income to be withdrawn from expenditure. In accordance with the necessary equality between assets and liabilities, what has been saved is lent by banks and spent by borrowers. In the case of macroeconomic saving, firms obtain part of the income formed by production either by the mechanism of profit or by borrowing it from wage earners. By definition, profit is that part of current income that is transferred to firms on the commodity market, to wit, that part of the income spent by income holders exceeding the cost of production of the goods and services purchased on the commodity market. The income thus transferred to firms is spent by them in the financing of a new production of capital goods. In other words, firms spend their profits on the labour market.

It should be clear that we are considering here only the case that leads to the formation of a macroeconomic saving, that is, uniquely what happens to that part of current income which will never be spent on consumption. It goes without saying that profits may be partly redistributed as interests or dividends; however, if they are, they will eventually be spent by creditors and shareholders on the commodity market. It is only when profits are invested that a macroeconomic saving is formed, and at this point it would be vain to look for any difference whatsoever between saving and investment. The same conclusion applies when firms borrow from savers their future profits and finance their current production of instrumental goods. Through the investment of profits (actually formed and advanced), part of current income is definitively transformed into fixed capital. The income thus forever withdrawn from consumption defines a macroeconomic saving, which is necessarily equal to the investment carried out by firms.

The reconciliation of the identity between S and I with the free will of savers and investors requires the introduction of a distinction between virtual and realized magnitudes. Before firms actually realize a profit or obtain a loan, their willingness or even their decision to invest are only desiderata, and investment itself is merely virtual. Income holders are free to take their decisions autonomously and can spend what they like (at least up to a certain point) on consumption. They can also freely decide how and to the benefit of whom they lend their savings. Income holders and firms thus jointly determine the amount of profit. However, before firms invest their profits (earned or 'borrowed'), macroeconomic saving remains a virtual magnitude. Firms and income holders' decisions may well diverge, and so will saving and investment as long as they are confined to the world of virtual magnitudes. Yet, once investment actually takes place it defines a realized magnitude, and so does macroeconomic saving.

What if savers had no intention to lend their saved income to firms? Would their decision create a gap between saving and investment? The answer is no, for the simple reason that savings are deposited with banks, and banks can but lend them. The reason is always the same: what is entered on the liabilities side of a bank's ledger is necessarily balanced by what is entered on its assets side. The income that savers do not lend to firms is lent by their banks, which is what financial intermediation is all about. As clearly stated by Keynes (1936/1946), savings do not disappear into the banking system, and are therefore not lost to investment. In the case we are examining here, firms are forced to purchase that part of output that is not purchased by consumers, and they

do so by spending the income saved by these very consumers and that banks lend to them. This forced investment pertains to the same category as the initial investment occurring at the moment production takes place: it is not a macroeconomic investment and does not determine a macroeconomic saving. However, if we consider the fact that firms necessarily spend the totality of the income that is not spent by consumers, it clearly results that saving and investment are but the two terms of an identity.

To conclude, let us criticize the concept of hoarding in this regard.

The idea that income can be saved in such a way as to take it out of circulation is closely related to a stock conception of money. If money were a stock endowed with a positive, intrinsic value, its circulation would resemble that of a physical fluid, say water or blood, within a water pipe or a system of blood vessels. Money could thus circulate more or less rapidly – and have a velocity of circulation as assumed by the quantity theory of money – or even stopped from circulating or lost through leakages, of which hoarding would be the most troubling example. Monetarism has greatly contributed to the spreading of this belief, and most economists still think it possible to hoard money in a way or another, by hiding banknotes under the mattress or into a safe. What these economists seem to miss is the fact that banknotes are claims on money, the most liquid claims available, definitely, but still claims. Money income lies behind these claims, and it cannot be hidden in any place whatsoever. As bankers know well, in order to get any given amount of banknotes their banks have to get indebted to the central bank, which means that part of the income initially deposited with commercial banks is indirectly entered on the liabilities side of the central bank's ledger, while income holders obtain, in the form of banknotes, a claim on the central bank. This amounts to saying that, even though banknotes can be hoarded, not even the smallest fraction of income can be lost to the economic system. The totality of income is deposited with the banking system, and, in compliance with the principle of double-entry bookkeeping, banks can but lend it.

The conception of money as a stock put into circulation by banks is both wrong and gravely misleading. It is wrong because it has nothing to do with the peculiar nature of economic exchanges, which, as shown by monetary macroeconomics, are absolute and not relative, and it is misleading because it nurtures the belief that economics can be understood by applying to it the conceptual tools of classical mechanics and mathematics. In fact, the development of economic analysis has been seriously hampered by the psychophysical approach advocated by Walras and his

followers. What characterizes economics and distinguishes it substantially from the other sciences is the nature of production considered as the instantaneous transaction by which physical output is given a numerical form. Bank money is a crucial element of this transaction. The economic laws discovered by Say, Keynes, and Schmitt acquire their deep significance and can be fully grasped only by referring to bank money and the principle that lies behind it.

This page intentionally left blank

Part II

Business Cycle and Crisis Theories: A Fundamental Critique

This page intentionally left blank

3

Business Cycles versus Boom-and-Bust Cycles

At the beginning of our discipline, economic crises were considered as disconnected and random events caused by exceptional and disparate contingencies and occurrences such as wars, political and social unrest, food shortage, natural disasters, and so on. Economists as well as bankers and businessmen soon started to realize that crises were much less isolated than it was generally thought, and that their recurrence was not random. Juglar (1857) was one of the first researchers to attempt to account for the recurrence of crises starting from statistical observation. Initially trained as a physician, the French economist is considered by Schumpeter (1954/1994: 1123) 'as to talent and command of scientific method, among the greatest economists of all times'. Starting from a series of statistical data about banknotes circulation, bank deposits, discount, and metallic reserves for France, England, and the United States – which he correlates to the volume of commercial transactions, the price of corn, rent, public revenue, and variations in population – Juglar (1857, 1889) asserts that monetary and commercial crises are strictly interrelated and are the unavoidable price to be paid for the evolution of capitalism.

Even if only implicitly, Juglar (1857, 1889) distinguishes between exogenous and endogenous causes, emphasizing the role played by the latter in generating a pathological state of affairs that is then exacerbated by external and occasional circumstances, which, by themselves, would never explain the regularity of economic crises. A contemporary of Juglar, Lawson (1848) had first distinguished between what he called the proximate cause of economic disruptions and their true cause, that is to say, 'between a cause common to all events and the occasional circumstances that actually give rise to the panic' (Besomi 2005: 26). As argued by Besomi (2005), Mills (1868) and Jevons (1884) shared

Juglar's and Lawson's point of view and looked also for the common causes capable 'to explain the remarkable appearance of regularity and periodicity which characterises these events [economic crises]' (Jevons 1884: 222). Some of these authors believed in a strict regularity in economic fluctuations (Jevons 1884), while others were less categorical with respect to cyclical periodicity (Juglar 1889). However, they all agreed that crises are not sudden disruptive events but the most dramatic phases of economic fluctuations that characterize our economic systems.

In the first section, we will analyse the search for a theory of crises where economic disturbances are considered as endogenous events inherent in the workings of our economic systems. Mainly developed by Marx (1839/1973), this approach is essentially macroeconomic and aims to determine the laws supposedly intrinsic to capitalism and inevitably leading to its self-destruction.

By contrast, the business cycle theories of crises, which we consider in the second section, aim to show that crises are periodical events due to economic fluctuations occurring in a world where economics is based on the principles of general equilibrium. Whether in the form of business cycle theories emphasizing the role played by trade, money, and credit or in the form of real business cycle theories, what the models proposed by Juglar's followers have in common is their microeconomic structure. The causes of economic disturbances and disorders are identified in exogenous shocks imputable to economic agents' (mis)behaviour.

This microeconomic approach also underpins Minsky's (1982, 1986) boom-and-bust cycle theories of financial crises considered in the third section. Despite its post-Keynesian origin, Minsky's analysis remains anchored in a microeconomic vision where the behaviour of speculators, banks, financial, and non-financial firms is considered as a potential cause of financial disturbances leading to financial bubbles and credit crunches.

Crises as endogenous disturbances

Marx is the author who best epitomizes the analysis of crises as endogenous events embedded in the very nature of capitalist economies. Although it may be claimed that Marx's analysis is too articulated to reduce his theory of crises to a simple model of pathological recurrences, it is not unjustified to emphasize the role of two concepts among many in Marxian thought, that is, the falling rate of profit and overproduction.

Let us start with what Marx (1839/1973) explicitly calls the most important of all economic laws, namely the falling rate of profit. According to Marx, profits have their origin in workers' labour-power and in the possibility for firms to purchase labour-power (a commodity) and use it in order to produce an output of a value greater than itself. The difference between the value obtained by putting labour-power at work and the remuneration of labour-power (real wages) is called surplus-value by Marx, and it is made up of goods produced free of cost and directly appropriated by firms. Now, the process of capital accumulation tends to concentrate an increasing amount of fixed capital (instrumental goods) in the hands of firms. This allows for an increase in physical productivity, which is not matched by a proportionate rise of surplus-value (because the latter is derived from labour and not from fixed capital). If we attend to the ratio between surplus-value (profit) and fixed capital, we find that its denominator tends to increase faster than its numerator, which is the reason why Marx considers the falling rate of profit as the unavoidable result of a process of capital accumulation that is marked by the inherent impossibility to reconcile the need for a growing fixed capital with the need to provide for its remuneration. If, as Marx maintains, the rise in surplus-value cannot make up for the rise in fixed capital, firms will be forced to decrease their investment in order to reduce capital accumulation, which will end up in a crisis of overproduction characterized by a sharp rise in unemployment.

In the view of the fact that the rate at which the total capital is valorized, i.e. the rate of profit, is the spur to capitalist production (in the same way as the valorization of capital is its sole purpose), a fall in this rate slows down the formation of new, independent capitals and thus appears as a threat to the development of the capitalist process; it promotes overproduction, speculation and crises, and leads to the existence of excess capital alongside a surplus population.

(Marx 1894/1981: 349–50)

Marx's other explanation of overproduction is also related to his theory of surplus-value. Being forced to respect the classical axiom that exchanges only take place between equivalents, Marx introduced a distinction between labour and labour-power in order to account for the formation of real profit. He was thereby able – or so he thought – to explain how firms can get hold of part of produced output without underpaying their workers. Wages pay for the exact value of what workers sell to firms, namely their labour-power. Yet, if wage earners are made

to work beyond the number of hours required to reproduce their labour-power, they provide firms with an amount of physical output obtained free of cost. The real goods thus obtained define Marx's surplus-value. A problem arises, however, when we pass from these real goods to the need to pay for them monetarily. To put it briefly, surplus-value goods have to be sold if firms are to earn a monetary profit.

The total mass of commodities, the total product, must be sold, both that portion which replaces constant and variable capital and that which represents surplus-value. If this does not happen, or happens only partly, or only at prices that are less than the price of production, then although the worker is certainly exploited, this exploitation is not realized as such for the capitalist and may even not involve any realization of the surplus-value extracted, or only a partial realization; indeed it may even mean a partial or complete loss of his capital.

(*ibid.*: 352)

The question is very clearly stated by Marx. The real goods making up for the surplus-value produced by workers must be sold on the market for produced output. If the totality of income available in the system is not enough to guarantee the final purchase of total output, firms suffer a loss and a crisis sets in.

What has to be carefully analysed at this point is whether Marx's conclusion derives from the contradictory nature of capitalism or from a logical inconsistency of Marx's own theory. Is capitalism such that it relies on a process of capital accumulation that inevitably leads to overproduction and that undermines itself since it leaves surplus-value goods unsold? Or is it not the case that the impossibility to sell profit goods results from a lack of consistency of Marx's analysis? Are surplus-value goods a shortcoming of capitalism itself, or is it a deficiency of Marx's theory that they cannot be realized monetarily? A rigorous analysis will show that it is the latter that is the case. In Marx's theory the explanation of profit is unsatisfactory, because it does not account for the monetary realization of surplus-value goods. The distinction between labour and labour-power is farfetched. The idea that labour-power is a commodity sold by workers to firms is not supported by facts and is nothing but a clever conceptual move to reconcile the exchange between equivalents with a plausible explanation of how firms can obtain for free a congruous part of produced output. Acceptance of this idea has the annoying consequence of leaving monetary profit unsatisfactorily explained, or even entirely unexplained. The only income that

the theory can account for is that earned by workers through the sale of their labour-power. However, the amount of wages is just enough to allow for the purchase of the real goods produced by workers as counterpart of their labour-power. Surplus-value goods are, by definition, given to firms free of cost, which means that they are excluded from the wage bill. Their production gives rise to no income in the first place so that no income is available for their purchase on the commodity market.

If Marx's theory were to provide a complete explanation of how profits are formed and realized, then it should attempt to show how and why the system functions in such a way that it inevitably leads to recurrent periods of overproduction. As Marx's theory is unable to explain how profit goods can be given their monetary expression, it is bound to leave us in the dark as for the origin and development of crises that ensue. The problem with overproduction is a fabrication of Marx's theory, and not a reflection of a pathological state of affairs in capitalism. Marx's theory is not inconsistent because the economic system is self-contradictory: it is Marx's own concept of labour-power that is responsible for the inconsistency of his theory, and it would be poor scientific behaviour to claim that his theory actually represents a state of affairs in the real world. A pathology in fact exists, but one needs a non-pathological theory to account for it and to provide a remedy eventually.

It is only fair to recognize that Marx was well aware of the realization problem, and that he tried to overcome it, unsuccessfully. Other authors, of whom Luxemburg (1913) is arguably the best example, attempted to find a solution consistent with the concept of surplus-value. Unfortunately, to refer to the takeover of new foreign markets as proposed by Luxemburg (1913) is of no avail, because it simply amounts to the enlargement of the analytical framework of reference, which does not alter the terms of the problem. Even if the income formed abroad could be spent in order to purchase the surplus-value goods produced domestically, the theory would have to explain how the goods produced abroad can be purchased once income is no longer available for that. If we consider the whole output (of all countries taken together) we immediately see that an amount equal to the sum of the surplus-value produced in the whole world cannot be realized: the income available is only sufficient to finance the purchase of a sum equal to that of the wage-goods of all countries.

Another attempt to solve, at least partially, the realization problem consists in claiming that banks can finance consumption through a generous credit policy (Roelandts 2009). This is a weird proposition indeed, as it amounts to saying that, in order to finance consumption, banks can

lend an income that does not yet exist. As bankers know very well, credit must be financed by deposits. If banks were to lend through an *ad hoc* creation of money, they would inject a number of empty money units into the system, which is clearly an inflationary operation that would not solve Marx's realization problem. What a theory must account for is the presence of an income that permits the monetary realization of surplus-value goods. A simple amount of money units created *ex nihilo* will not do, because they will be endowed with no purchasing power at all.

The failure of Marx's theory to solve the realization problem sets the limit of his own explanation of economic crises. Indeed, overproduction is not due to Marx's mechanism of surplus-value formation, and the fall of the rate of profit cannot be explained by referring to Marx's distinction between labour and labour-power. As a consequence, Marx's claim that economic crises may recur every ten years, when caused by the need to renew fixed capital, or every 30 years, when due to the fluctuation of the rate of profit, becomes irrelevant. Likewise, his prediction of a culminating crisis leading to the collapse of the capitalist system is not theoretically supported by a consistent analysis of capital accumulation. Yet, as we shall notice in Chapter 6, despite Marx's failure to provide a sound analytical foundation to his own theory, his intuitions about the fall of the rate of profit and the limits to capital over-accumulation are prescient, and find a correct explanation and their *raison d'être* in Schmitt's approach to monetary macroeconomics.

The passage to business cycle theories

The works of Lawson (1848), Juglar (1857, 1889), Mills (1868), and Jevons (1884) have greatly contributed to the birth of business cycle theories. In 1913 Mitchell was thus able to claim without any serious risk of being contradicted that '[c]rises are no longer treated as sudden catastrophes which interrupt the "normal" course of business, as episodes which can be understood without investigation of the intervening years. On the contrary, the crisis is regarded as but the most dramatic and briefest of the three phases of a business cycle' (Mitchell 1913: 5). The recurrence of crises was widely accepted as a fact, and was ascribed to endogenous causes, inherent in the process of economic growth. It is symptomatic, in this respect, that Schumpeter (1950: 149) praises Juglar, 'the outsider who may be said to have created modern business cycle analysis', for his 'all-important [...] diagnosis of the nature of depression, which he expressed with epigrammatic force

in the famous sentence: the only cause of depression is prosperity' (Schumpeter 1954/1994: 1124). Schumpeter (*ibid.*: 1135) was deeply convinced that economic fluctuations are embedded in the capitalist system, and that it is necessary 'to look to business cycles for material with which to build the fundamental theory of capitalist reality'.

Like Schumpeter, a great number of economists embracing business cycle analysis advocate the theoretical framework of general equilibrium. However, general equilibrium seems at odds with the idea of cyclical fluctuations. According to neoclassical analysis, market forces lead towards equilibrium, whereas business cycle theories investigate a situation of constantly renewed disequilibrium. It is true that equilibrium does not need to be static, and economic growth theories are meant to deal with moving equilibrium. But business cycle fluctuations are unlikely to be a component of economic growth theories proper: 'it has been common to think of business cycle models as separate from models of economic growth and to characterize business cycles as the deviations from some smooth, usually deterministic, trend that proxies for growth. Theories of the business cycle are then constructed to explain these deviations' (Plosser 1989: 59). Whether or not we agree with Hicks's (1965: 4) claim that 'the distinction between trend and fluctuation is a statistical distinction', it is clear that slumps can hardly be considered as states of negative economic growth consistent with a moving trend of the economy in line with the principles of general equilibrium analysis.

This is not a unanimously shared point of view. Prescott (1999), for example, maintains that economic growth theory may prove useful in explaining business cycle fluctuations. 'The developers of growth theory thought the theory would be useful for studying long-term growth issues but that a fundamentally different theory would be needed for studying business cycle fluctuations. Once the implications of growth theory were derived, however, business cycle fluctuations turned out to be what the theory predicts' (Prescott 1999: 26). Prescott's claim rests on empirical observations but also on the theoretical claim that '[e]ssentially, *business cycles* are responses to persistent changes, or *shocks*' (*ibid.*: 27). According to Prescott, a shock shifts the constant growth path down and 'the economy converges to its new lower constant growth path' (*ibid.*: 27). But what about long-period slumps? When economic crises last for a long time, how do downward fluctuations relate to a hypothetical constant growth rate? And, even more important, how can the concept of equilibrium be substantiated when the economy is constantly submitted to up and down fluctuations, which, *de facto*, define a situation of moving

disequilibrium? Neoclassical economists have attempted to answer these questions mainly in two different ways. The majority of them has followed Hayek in looking for the role played by monetary shocks, while others have stuck to real exchanges and developed what is known as real business cycle analysis. Let us elaborate on this.

Hayek's starting point is the logical impossibility to reconcile static general equilibrium analysis with trade cycle theories. According to the Austrian economist,

[t]he obvious, and (to my mind) the only possible way out of this dilemma, is to explain the difference between the course of events described by static theory (which only permits movements towards an equilibrium, and which is deduced by directly contrasting the supply of and the demand for goods) and the actual course of events, by the fact that, with the introduction of money (or strictly speaking with the introduction of indirect exchange), a new determining cause is introduced.

(Hayek 1933/2008: 6)

The analyses developed by von Mises (1928), Böhm-Bawerk (1884/1959), Wicksell (1898/1965), and Friedman (1970) are among the best attempts to account for the role played by money without giving up the main principles of neoclassical analysis, and many others have followed in their footsteps. Have they been successful? To answer this question we will have to recall our analysis about the nature of money and the way it is associated to output. For the time being, let us simply observe that the increasing importance attributed to money has led to the spread of credit cycle analysis, which has become a key feature of the majority of theories aiming to explain economic fluctuations.

At the other end of the spectrum we find those economists who, following Hicks's (1933/1982: 32) claim that the 'idealised state of dynamic equilibrium [...] give[s] us a way of assessing the extent or degree of disequilibrium', rely on the role played by real disturbances in order to explain business cycle fluctuations. Their idea is that economic agents behave rationally (Muth 1961, Lucas 1972), and that business cycles can be explained by referring to the way agents respond to relative price movements. In general, '[r]eal business cycle models view aggregate economic variables as the outcomes of the decisions made by many individual agents acting to maximize their utility subject to production possibilities and resource constraints. As such, the models have an explicit and firm foundation in microeconomics' (Plosser 1989: 53). To put it briefly, real business cycle models are meant to represent

'a fully articulated artificial economy which behaves through time so as to imitate closely the time series behaviour of actual economies' (Lucas 1981: 219), and account for economic fluctuations caused by technology, productivity disturbances, and 'other shocks arising from preferences, government, [and] terms of trade' (ibid.: 71).

Can business cycles be derived from statistical observation?

Let us go back to Juglar (1889), and to his attempt to construe an objective theory of business cycles starting from a collection of statistical data. The fact that Juglar was not a well-trained statistician is of no importance here. Our question goes beyond technicalities, and cannot be answered by referring to current or future developments in the field of statistics. Our concern is the conceptual possibility to derive 'the law of crises and their periodicity [...] from factual observation only' (Juglar 1889: xv, our translation). As emphasized by Schumpeter (1954/1994: 1124), Juglar 'proudly claimed to have discovered the "law of crises" without any preconceived theory or hypothesis'. Was he right? Can economic laws be derived directly from statistical time series? Like Juglar, many other economists since answer yes, and are deeply convinced that empirical observation is the appropriate starting point for the construction of a theory that is not a dogmatic interpretation of reality. Statistical data are what they are – so we are told – and if a correlation proves to exist between them, we are entitled to infer a causal link without reverting to any preconceived notion or hypotheses. If this were indeed the case, statistical data collection and elaboration alone could and should replace economic analysis, and we could reasonably expect a general and objective theory to result from empirical observation.

One of the authors who most clearly reject the centrality of the statistical method is Hayek. In his 1933 article on trade cycles he claims that

[e]mpirical studies, whether they are undertaken with such practical aims in view, or whether they are confined merely to the amplification with the aid of special statistical devices of our knowledge of the course of particular phases of trade fluctuations, can at best afford merely a verification of existing theories; they cannot in themselves provide new insight into the causes or the necessity of the trade cycle.

(Hayek 1933/2008: 3)

The Austrian economist's denial to endorse the priority of statistics with respect to economic analysis rests on two considerations: the

impossibility to collect and elaborate data without the help of a theory, and the impossibility to derive a law from direct observation of time series data. Let us analyse these two arguments in some detail.

The following quotation sets the first point with great clarity.

The dependence of statistical research on pre-existing theoretical explanation hardly needs further emphasis. This holds good not only as regards the practical utilization of its results, but also in the course of its working, in which it must look to theory for guidance in selecting and delimiting the phenomena to be investigated. The oft-repeated assertion that statistical examination of the trade cycle should be undertaken without any theoretical prejudice is therefore always based on self-deception.

(*ibid.*: 5)

This amounts to saying that economic reality is not directly given or unambiguously offered for us to observe and represent as it is. Even in physics, factual observation may be deceptive, and Gould (2007: 291) has no hesitation in claiming that, in the field of natural sciences, '[f]acts do not "speak for themselves"; they are read in the light of theory'. Economics is no exception. Moreover, economic magnitudes are not intrinsic and observable qualities of things. A great number of economic 'objects' are but conceptual constructs. Furthermore, even when an economic magnitude refers to a concrete object, it does not pick out any of its physical features. A price is not a material property of produced output, and it does not manifest itself to an objective observer, who can straightforwardly compute it in order to relate it to other data observed via a process of direct perception. As unanimously recognized, the price of produced goods and services is a concept and, as such, can only be defined and measured by means of a theory. This means that, from the outset, the collection of statistical data is biased by the choice of the theory operated by the statistician. It is the chosen theory that tells us what price, income, capital, and so on are meant to signify, and how they have to be measured. Applied to business cycle analysis this amounts to saying that a theory is required not only to explain how economic cycles occur, but, primarily, what they are. 'There seems to be no way to determine how business cycles are to be dealt with short of understanding what they are and how they occur' (Lucas 1981: 235).

As stated by Hayek (1933/2008), the asserted objectivity of statistical analysis of trade cycles is self-deceiving, statisticians being forced to choose a theory that influences their whole work. Indeed, a theory is

a consistent set of interconnected concepts that introduces a series of conceptual definitions deduced from a set of principles it posits. Once a theory has been chosen, whether explicitly or implicitly, it affects the way statistical data are collected and elaborated as well as the way they are compared to one another in search of a hypothetical correlation. The true problem is therefore not that of construing an objective statistical model supposedly derived from direct factual observation, but finding a theory able to explain the very nature of economic phenomena and their logical interrelations without resting on any dogmatic assumption, whether stated explicitly or surreptitiously introduced in the form of axioms.

Hayek's epistemic assessment as to the possibility of statistics to establish any logical interconnection between economic magnitudes is very clearly stated. In particular, he argues that 'it must be clearly recognized that the use of statistics can never consists in a deepening of our theoretical insight' (Hayek 1933/2008: 4), and advocates the priority of logically determined interrelations over statistical correlations. 'Empirically established relations between various economic phenomena continue to present a problem to theory until the necessity for their interconnections is demonstrated independently of any statistical evidence' (ibid.: 4). To put it simply, the problem is as follows: how can statistical data observation reveal the existence of a causal relationship between economic events? Logic has us to answer that it cannot. Even if we were to observe for a very long time that event B always follows event A, we could not infer with certainty that A is the cause of B, because it would be enough for once B not to follow A for our conjecture to be proven false. In practice, things are much more complex, and it is self-deceiving to isolate two events from all the others in order to show that they are causally interrelated. The possible correlations are too numerous, and there are too many cross-references among economic events and magnitudes to rely on statistical methods to determine or specify economic laws. This difficulty cannot be overcome even by the most sophisticated techniques once it is realized that data collected at different periods of time are likely to be heterogeneous, which makes it impossible to compare them directly, and that, in the absence of a theory capable to determine a set of laws, the number of combinations of constantly changing data, which cannot be viewed as homogeneous, becomes infinite.

Statistical methods and econometrics are nevertheless widely used in business cycle analysis, in search of significant or robust correlations that may serve as the basis for an explanation of economic fluctuations.

Thus, for example, Reinhart and Rogoff (2008, 2009) develop a historical analysis of financial crises starting from time series data referring to standard indicators 'such as asset price inflation, rising leverage, large sustained current account deficits, and a slowing trajectory of economic growth' (Reinhart and Rogoff 2009: 466). Through the simple observation of historical data, they feel confident in concluding that '[a]n examination of the aftermath of severe financial crises shows deep and lasting effects on asset prices, output, and employment' (ibid.: 471). Apparently, Reinhart and Rogoff's analysis is straightforward, and its conclusions can be shared by anyone, as the authors merely describe the events marking and accompanying financial crises, without attempting to establish any causal relationship between them. To the extent that Reinhart and Rogoff's work is a mere description of what happens during a financial crisis, it may prove useful in helping to support or oppose a given theory, but it cannot aim at providing any causal explanation by itself. Yet, as the title of Reinhart and Rogoff's 2009 article suggests, its aim goes beyond a simple description, its authors' ambition being that of determining the consequences (aftermath) of financial crises. In this respect, factual observation lacks scientific rigour, and is absolutely insufficient to provide the basis for a causal interpretation. The point at stake is whether or not the concomitance or the regular succession of events as described by statistical data allows to derive any necessary relationship between them. To observe that financial crises are accompanied by a decline in prices, output, and employment is not enough to conclude that the two sets of events are causally connected. Analysis could well prove, in fact, that financial crises are an effect rather than the cause of economic crises, and thus discredit any thesis based on factual observation.

This is not to say that statistical methods have no relevance at all in economic analysis. In fact, a distinction should be made between micro- and macroeconomic analyses of crises.

From a micro to a macroeconomic analysis of crises

As far as macroeconomics is concerned, it is hopeless to rely on statistics to determine the laws governing the economic system as a whole. Macroeconomic concepts can neither be defined nor measured by means of statistics. However, once these concepts are specified, they allow us to make sense of microeconomic variables, whose fluctuations can then be statistically reproduced and correlated. An example may prove useful to illustrate this point. The price of produced output is first and foremost a macroeconomic concept defined by the relationship established by production between output and money. Statistics is of no

use in defining it, and provides no clue on how it has to be measured. Yet, once the macroeconomic price is known, the (microeconomic) selling price fixed by firms acquires a precise meaning and its evolution can be followed through time. Empirical data concerning microeconomic prices can be collected and confronted, for instance, with data relative to (microeconomic) demand, and an attempt can be made to correlate the two series of data. In general, the statistical method may be applied to microeconomics, but is unable to support or justify the macroeconomic framework of analysis required for the understanding of an actually existing economic system.

The reader well acquainted with mainstream economics will probably be surprised by, and vehemently disagree with the preceding statement, confident as s/he is that macroeconomics can be built or based solely on microeconomic foundations. It is thus convenient to elaborate on the relationship between micro- and macroeconomics, and consider how these two approaches are currently integrated into business cycle theories.

As Lucas (1977) maintains, the belief that macroeconomics must be founded on microeconomics was already commonplace in the 1970s. 'The observations that macroeconomics is in need of a microeconomic foundation has become commonplace' (Lucas 1977: 8). Contrary to what is generally believed by most economists, however, we maintain that macroeconomics cannot be derived from microeconomics. The search for microeconomic foundations of macroeconomics is a self-defeating quest, because macroeconomics both conceptually and institutionally precedes microeconomics. If economic laws could not be instituted at the macroeconomic level, microeconomics could not even exist. This is a very strong and peremptory statement, which has to be substantiated by rigorous and objective analysis. Quantum macroeconomics (see Schmitt 1984a, Cencini 2001, Rossi 2001) provides the necessary proofs in this respect. There is no need to elaborate on them at this juncture, as it is enough to emphasize some of the main results of quantum macroeconomics. In particular, it is worth pointing out that Schmitt (1996a) gives a final proof that relative (microeconomic) prices are logically undetermined, unless they are derived from absolute (macroeconomic) prices. Now, absolute prices result from the association between money and output, an event that is not influenced by agents' forms of behaviour. Unlike microeconomics, macroeconomics does not attempt to describe how economic agents behave in order to maximize their utility, but rather explains the logical, objective framework presupposed by the system within which economic agents take their decisions. By proving that prices cannot

be determined through direct (relative) exchange, quantum analysis shows that goods and services cannot be made homogeneous through subjective evaluation. This clearly means that either economic analysis is based on macroeconomic foundations, or it is bound to fail for lack of a proper object of inquiry. Goods and services are the proper objects of economic analysis if, and only if, they can be measured by a common, a-dimensional standard, that is, through the determination of their money price, an objective that can be attained only if the dogma of macroeconomics being founded on microeconomics is jettisoned.

As this book is meant to present the main elements of a truly macroeconomic analysis of crises, it should be noted at this stage that business cycle analysis has traditionally followed mainstream economics in looking for a microeconomic representation and explanation of cyclical fluctuations. This is not surprising, of course, since business cycle theories are nothing but general economic theories applied to the study of economic disturbances or disorders. To illustrate this point let us just consider a few examples.

In his analysis of what he calls 'the return of depression economics', Krugman (1999) claims that the liquidity trap played a decisive role in Japan's crisis of the late 1990s, and explains it by referring to Japan's demography and to the decisions taken by consumers and firms. 'Japanese consumers are saving for retirement, even while firms are unwilling to invest, because they expect a shrinking market' (Krugman 1999: 68). Insofar as he identifies the cause of disturbances in the behaviour of consumers and firms, Krugman endorses a microeconomic approach, thus sharing the neoclassical method of his neoclassical colleagues. Diamond and Rajan (2001) follow a similar path when they point to the behaviour of banks as the main cause of financial crises. 'As we show, a significant, albeit temporary, postponement of cash flows is enough to bring down a banking system that has financed itself with short-term debt. There is no doubt, therefore, that a banking system that exhausts its short-term debt capacity renders itself more prone to crisis' (Diamond and Rajan 2001: 39). Even Bernanke (1995), who explicitly investigates the macroeconomics of the Great Depression, is not immune from founding his analysis on microeconomic considerations. As the following quotation shows, he refers to borrowers', firms' and households' behaviour as causal factors inducing financial distress.

To the extent that potential borrowers have unique or lower-cost access to particular investment projects or spending opportunities,

the loss of borrower net worth effectively cuts off these opportunities from the economy. Thus, for example, a financially distressed firm may not be able to obtain working capital necessary to expand production, or to fund a project that would be viable under better financial conditions. Similarly, a household whose current nominal income has fallen relative to its debts may be barred from purchasing a new home, even though purchase is justified in a permanent-income sense. By inducing financial distress in borrower firms and households, debt-deflation can have real effects on the economy.

(Bernanke 1995: 18)

Numerous other examples could be given, which illustrate the microeconomic approach to crises shared by neoclassical economists as well as by New or post-Keynesian scholars. They explain indeed the different phases of business cycles by emphasizing the role played by expectations, self-interest, maximization, and other such factors influencing economic agents' decisions about consumption, saving, investment, credit, and so on. They thus define recovery as a phase in which profits increase, business expectations improve, and optimism prevails so that investment is boosted and credit expands. Prosperity is then to follow, with larger profit expected as well as increases in investment and consumption and a rise in prices. In this approach recession is described and explained as the decrease in business expectations and consumption. Optimism gives thereby way to pessimism, production is reduced, and credit contracted. Finally, if depression sets in, consumption and investment reach a minimum, credit shrinks, and interest rates fall until once again expectations improve, and optimism returns.

To be true, this simplified account of business cycle fluctuations is far from being exhaustive. It is good enough, however, in showing how deeply economic analysis is anchored in microeconomics. We claim instead that a correct understanding of crises requires a complete shift in the baseline of the analysis, which means switching from a micro to a macroeconomic analysis. Without this radical shift in the starting point of the analysis, economics is bound to remain bogged down in the hopeless search for mathematical-computational models able to simulate a reality that is far too complex to be captured thereby.

The conceptual shift we advocate will sound at odds with the prevailing view that is considered common sense. It is undisputable, for example, that a single individual, a firm, or a State may get too much indebted, and even run into bankruptcy. When this happens, serious consequences can ensue, not only for the individual, the firm or the

State, but also for those who are financially involved with them. Bad management, undoubtedly, leads to disturbances that may well affect a substantial number of economic agents. Is it not correct to speak in this case of an economic crisis? Most economists would probably not hesitate to answer yes to this question, for the simple reason that they are accustomed to reason in microeconomic terms, and to locate the causes of economic fluctuations in the behaviour of economic agents. They would be right if economic crises were not defined as pathological situations affecting the economic system as a whole. Our key objection here is that mismanagement does not by itself transform a sound economic system into a disorderly one, and can therefore never be the origin of economic crises. A series of wrong decisions is often likely to lead to important losses for the people involved, but it does not jeopardize the entire system. In other words, economic crises are macroeconomic events whose origin cannot be of a microeconomic nature. The great economists of the past have not been fooled by appearances, and have consistently looked for causes altering the system as a whole, and not simply affecting some of its elements, which they considered as symptomatic of underlying factors. We share their point of view, and it will be our contention in this book that crises are due to the misrecognition of macroeconomic factors or constraints, and the pathological process of capital accumulation that ensues.

Despite the enormous technical difficulties owing to the essentially unforeseeable behaviour of individuals, one could still hope to develop more realistic models if economics was uncontroversially viewed as a branch of mathematics. This would be the case if it could be possible to maintain that goods and services are unambiguously number-like. This is far from being so, let alone proven. Indeed, is there any need to show how absurd such a claim is? The rhetorical character of this question should be clear, and it should also be obvious that scientific analysis can only discard as metaphysical the assumption that goods and services are essentially numbers. Actually, numbers are all important in economics; however, their introduction cannot be treated as self-evident, or assumed as given a priori. Debreu's (1959: 30) claims that '[a] quantity of wheat can be any real number' or that '[a] quantity of well-defined trucks is an integer' (ibid.: 30) beg the question. The physical measure of an object is not a pure number, but a number associated with a physical dimension. A quantity of wheat is measured in units of a dimensional standard such as kilos, tons, and so on, and it would be nonsensical to maintain that 100 tons of wheat are identical to the number 100. By way of his magical device, which is no other than

the mathematical formalism he deploys, Debreu (1959) transforms 100 tons in the number 100, and is thereby able to consider economics as mathematical in nature or, as he maintains, axiomatical. In actuality, the statement that a dimensional quantity is nothing but a number is far from being a self-evident truth, let alone a proven one.

If economics is to be a science, its object of investigation must be defined and measured; no one disputes this. Following Walras (1874/1984), Debreu (1959) correctly maintains that goods and services are measured in numbers (*numéraire*), but, like his predecessor, he fails to explain how physical goods and services can be transformed into numbers, that is, how numbers can be integrated with produced output, to wit, how numbers can become the numerical form of products. Now, the introduction of numbers into economics is not a mathematical problem. Mathematics does neither explain how this integration occurs, nor what its deep meaning is. While mathematics deals with pure numbers, that is to say, with purely numerical entities, economics is concerned with output and the transactions that, by associating it to numbers, transform a heap of physically heterogeneous objects into a set of homogeneous goods. This task cannot be fulfilled by mathematics.

Economic principles or laws can only be derived from an analysis that is not mathematical in conception. This corollary follows directly from the logical impossibility to identify physical objects as numbers, and it admits of no exception. No wonder, then, that the dominant mathematical modelling of economic reality has failed to attain its ambitious objectives. Despite the use of ever more sophisticated formal analytical tools, econometricians and statisticians are struggling in vain to fit economic reality to a set of mathematical rules and devices. Their results are as elegant as they are deceiving: thousands of extremely complex models tell us precious little about the state of real world economies. The nature of crises, for example, is still as mysterious as it was when economists started to investigate economic fluctuations. Explanations have grown in number, and possible factors of economic disruption have increased along with their hypothetical interrelations. As is often the case, this is symptomatic of the intrinsic inadequacy of the available analytical tools. In their search for a better understanding of crises, economists have been misled by the application of mathematics and statistics, and by the related and erroneous belief that economics needs to be based on microeconomic foundations.

The different sorts of crises that can be envisaged are too numerous and intertwined to be presented in a significant and exhaustive way. However, a broad albeit meaningful distinction is worth being

introduced, namely between business cycle theories emphasizing the role played by money and credit, and real business cycle theories. Let us briefly examine how this distinction has influenced traditional business cycle analysis.

Credit cycles versus real business cycles

The idea that fluctuations in credit may be the cause of business cycles was already present in Juglar's (1857) work. The French economist claimed that credit expansion is what creates the necessary condition for business cycles to occur. His analysis was essentially microeconomic, and rested on the assumption that speculation has a disruptive impact on production and exchange. 'We can interpret Juglar's theory as based on speculative behaviour but also on a contagion phenomenon acting, by definition, at a collective level. [...] The different and inter-related phases of the cycle can be seen as the outcome of a pure credit adjustment or as permanent oscillations of the system between confidence, credulity and then distrust' (Dal-Pont Legrand and Hagemann 2007: 17).

The author who most clearly singles out bank credit as the origin of business cycle fluctuations is Hayek (1933/2008), who uncompromisingly argues that '[s]o long as we make use of bank credit as a means of furthering economic development we shall have to put up with the resulting trade cycles' (Hayek 1933/2008: 31). What Hayek had in mind was essentially the idea that, through the mechanism of depositing and lending, bank credit may be multiplied and extended beyond the needs of economic activity. Economists have since gone much further in their analysis of the effect of financial disruptions, and the financial crisis that burst in 2007 seems to have left no doubt about the influence of credit on economic activity. As undisputable as this causal link might appear, it is nevertheless far from being completely and properly understood. Hence, it is still unclear whether the problem lies with money, credit, or speculation.

If we consider the analyses of Friedman (1956, 1970), Lucas (1977), Bernanke (1995), and many others, we are led to conclude that money and monetary policy are the key components of external disturbances triggering business cycle fluctuations. Friedman (1956, 1970) is notably the author who has contributed the most to the reappraisal of the quantity theory of money.

The viewpoint of pure neoclassical economists differs substantially from Friedman and his followers' inasmuch as its advocates are convinced that economic fluctuations have their origin mostly in

disturbances affecting real variables. 'The RBC [real business cycle] models of the 1980s [...] imply no scope at all for monetary stabilization policy because real variables are modelled as evolving in complete independence of any nominal variables' (Woodford 2003: 6). Real business cycle theories are built on the belief that economic agents' decisions are the true cause of economic fluctuations, and that they are affected by changes in technology, productivity, government policy, preferences, terms of trade, or even money (if only exceptionally). The microeconomic nature of real business cycle theories is clear, as is their strict dependence on the concept of equilibrium.

Real business cycle models view aggregate economic variables as the outcomes of the decisions made by many individual agents acting to maximize their utility subject to production possibilities and resource constraints. As such, the models have an explicit and firm foundation in microeconomics. More explicitly, real business cycle models ask the question: How do rational maximizing individuals respond over time to changes in the economic environment and what implications do those responses have for the equilibrium outcomes of aggregate variables?

(Plosser 1989: 53)

Our earlier criticism regarding microeconomics and its suitability for providing an appropriate framework of analysis for the study of crises apply here. In particular, the logical indeterminacy of relative prices (Schmitt 1996a, 1999b) leads to a dramatic conclusion: no economic theory can be built logically on direct (relative) exchange. Further, it is beyond doubt that all contemporary economies are monetary, which makes it meaningless to reason 'as if' they operated in a dichotomous world. The presence of money cannot be disputed, even though its ubiquitous nature remains enigmatic. Both the need for money and the conceptual impossibility of identifying it with a commodity follow from the indeterminacy of relative prices. It thus appears that neoclassical analysis is at odds with actually existing and functioning economies. Any attempt to explain their working in 'real' terms is therefore doomed to failure. Real business cycle theories can be no exception: their microeconomic foundations cannot support the building of a theory consistent with the monetary nature of today's economies. In this respect, it is worth observing that a critical adherent to neoclassical theory like Hayek had nevertheless clearly intuited that a theory of crises, or business cycles, must put money on centre stage: 'only the assumption

of primary monetary changes can fulfil the fundamentally necessary condition of any theoretical explanation of cyclical fluctuations – a condition not fulfilled by any theory based exclusively on “real” processes’ (Hayek 1933/2008: 15).

What we have been examining so far are but a few introductory considerations about the shortcomings of mainstream analysis. Crises are still to be explained. As noted by Böhm-Bawerk (1898: 132, quoted by Dal-Pont Legrand and Hagemann 2007: 9), ‘[a] crisis theory can never be the inquiry of a separate part of socio-economic phenomena but it is, if it should not be a dilettante absurdity, always the last or last but one chapter of a written or unwritten socioeconomic system, the mature fruit of the perception of the complete socioeconomic phenomena and their interdependence’. The Austrian economist has expressed it well: crises can only be explained if we are able to compare a pathological or dysfunctional situation to a healthy state. For such comparative assessment a deeper understanding of the laws governing the working of the whole economic system is necessary. This is why the positive reconstruction we offer is based on the analysis of the macroeconomic foundations of macroeconomics, where monetary macroeconomics is essential to explain the orderly and the disorderly working of our capitalist economies.

In fact, a tentative approach to monetary macroeconomics has been proposed by those researchers, working in the tradition of either Austrian or post-Keynesian economics, who endeavour to explain crises with regard to the evolution of the business cycle as a result of financial actors that operate within the economic system. Let us consider this approach in the remainder of this chapter.

Boom-and-bust cycles as endogenous phenomena

A heterodox analysis of crises is offered by those economists who, following Minsky (1982, 1986), point out the essentially unstable nature of our capitalist systems based on financial motives, financial markets, financial products, and financial institutions (see Epstein 2005; Wray 2009, among others). This strand of thought, which is often ranged into post-Keynesian economics (see Chapter 5), elaborates on Keynes’s representation of ‘financial circulation’ as being separated from (and dominating on) ‘industrial circulation’ (Keynes 1930/1971: 217). Distinguishing ‘the deposits used for the purposes of industry, which we shall call the *industrial circulation*, and those used for the purposes of finance, which we shall call the *financial circulation*’ (ibid.: 217), Keynes

pointed out that ‘the volume of trading in financial instruments, i.e. the *activity* of financial business, is not only highly variable but has no close connection with the volume of output whether of capital goods or of consumption goods’ (ibid.: 222). According to Keynes (1930/1971: 227), ‘changes in the financial situation are capable of causing changes in the value of money in two ways. They have the effect of altering the quantity of money available for the industrial circulation; and they may have the effect of altering the attractiveness of investment.’ To be sure, Keynes views investment as a macroeconomic operation that generates income in the whole system, because it originates newly produced output in what he calls the industrial circulation or ‘enterprise’. By contrast, he defines speculation as ‘the activity of forecasting the psychology of the market’ where financial assets are exchanged (Keynes 1936/1946: 158), in order to capture a rent derived from produced income. To put it in Keynes’s own words:

Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation. When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done.

(ibid.: 159)

Minsky (1975, 1977) elaborated on this view, to explain his now-famous ‘financial instability hypothesis’, that is, ‘a theory of how a capitalist economy endogenously generates a financial structure which is susceptible to financial crises, and how the normal functioning of financial markets in the resulting boom economy will trigger a financial crisis’ (Minsky 1977: 15).

Contrary to Keynes (1936/1946), whose analysis focuses on the factors determining firms’ investment, Minsky (1977, 1982, 1986) investigates how the different means of financing investment are exploited by both financial and non-financial firms, and how this generates what he labelled ‘financial fragility’, a process leading eventually to financial instability and crisis. Minsky’s theory, which has influenced the so-called ‘Bank for International Settlements’ macro-prudential approach’ to financial stability (see Bank for International Settlements 2001, Borio and Lowe 2002, Borio and White 2004), provides a microeconomically founded explanation of a macroeconomic issue such as financial crises. It thus contributes to (rather than criticizing) the evolution of the orthodox methodological approach to macroeconomics, although

it establishes the microeconomic foundations of macroeconomics on post-Keynesian (rather than on neoclassical) grounds.

Contrary to neoclassical thinking, Minsky argues that crises are not the result of some exogenous factors but originate in specific endogenously operating mechanisms of a monetary production economy. He notably points out a 'paradox of tranquillity' – or that 'stability is destabilizing' – arguing that a protracted period of economic stability induces firms and their creditors (particularly banks) to increase their leverage, and in particular the firms' debt, beyond sustainable levels, inflating a credit or asset bubble that bursts when agents' expectations about future profits change considerably and for a relevant number of actors (Minsky 1975). In particular, according to Minsky (1982: 86), '[b]anking, that is, the financing of capital asset ownership and investment, is the critical destabilizing phenomenon'. Considering the loans-create-deposits mechanism that characterizes endogenous-money theory, Minsky argues that banks are willing to grant further loans to already-indebted firms, when both categories of agents expect that the future profitability of these firms will enable them to serve their debt out of expected profits: during the expansionary phase of the business cycle, these firms become thus more and more indebted (as a percentage of both their assets and their income), thereby making the whole financial system (most particularly its banking institutions) increasingly fragile as time goes by. This 'systemic' financial fragility is reinforced by financial innovations that banking institutions exploit in order for them to increase both their profitability and market shares in an attempt to avoid regulatory constraints (Minsky 1992: 17). The boom-and-bust cycle, which ends with a financial crisis as a general rule, is therefore the result of the behaviour of financial and non-financial firms whose interconnectedness across a variety of markets and countries has a number of macroeconomic consequences during the whole business cycle. This is so much so when financial institutions provide credit also to the household sector (beyond the business and general government sectors): the 'subprime' meltdown observed in 2006–07 in the United States has indeed been explained via Minsky's 'financial instability hypothesis' by the advocates of the so-called 'modern money theory' spearheaded by Wray (2008, 2012). 'When losses on subprimes began to exceed expectations that had been based on historical experience, prices of securities began to fall' (Wray 2008: 53). Indeed, as already explained by Minsky (1957a), financial innovations induce a reduction of the ratio between liquidity and total assets held by financial institutions – where liquidity is understood as 'the amount of

money which is held as a liquid asset' (Minsky 1957b: 864) – thereby increasing the ratio between the financial institution's debt and its net worth. When this becomes plain in a framework where profit expectations turn out to be much exaggerated with respect to actual earnings and an increasing number of financial obligations are not met by their debtors, then leveraged financial (and non-financial) institutions must dispose of an increasing amount of illiquid assets, in a desperate as well as vain attempt at recovering enough liquidity to pay their own creditors – who are successively taken in a 'debt-deflation' spiral already depicted by Fisher (1933: 344) when he observed that *'the very effort of individuals to lessen their burden of debts increases it, because of the mass effect of the stampede to liquidate in swelling each dollar owed'*. Indeed, as noted by Wray (2008: 57), 'as the subprime market unravels, fears spread to other asset-backed securities, including commercial real estate loans, and to other bond markets such as that for municipal bonds. Markets are beginning to recognize that there are systemic problems'.

The Minsky explanation of boom-and-bust cycles is intriguing and close to empirical evidence, particularly as regards the global financial crisis that erupted after the credit and housing bubbles burst in the United States in 2007. This explanation, however, is partial, as it focuses on the behavioural factors of that crisis (and of financial crises in general). To be true, some scholars working in the footsteps of Minsky have proposed to reform the financial system, in order to avert that some forms of behaviour by both banks and non-bank financial institutions could affect the whole economy negatively (see, for instance, Kregel 2009, Nersisyan and Wray 2010, Wray 2010). Nevertheless, these Minskyan reform proposals address financial-behavioural rather than monetary-structural factors of financial crises. They are therefore not enough to avert that further 'systemic' crises happen again, since it is the system as a whole (and not only its agents) that does not work properly – to wit, according to the macroeconomic laws of a monetary production economy. These laws cannot be seized and understood via a microeconomically founded analysis of macroeconomic issues such as a financial and economic crisis. Indeed, the solution to eradicate the roots of a systemic crisis cannot be limited to avoid agents' (mis)behaviour that affects the working of the economy as a whole; it must affect the working of the whole system – hence go beyond behaviour – considering the Aristotelian principle that the whole is different than the sum of its parts and that macroeconomics is much more than the mere aggregation of microeconomic events.

What is missing in Minsky's analysis is an explanation of the pathology at the origin of the capital that economic agents use to finance their speculative transactions. The idea that such a capital may result directly from the credit creation operated by financial institutions is naïve, not to say inconsistent with the principles of double-entry bookkeeping followed by these institutions everywhere in the world. In reality, the mechanism leading to the formation of pathological capital is far more complex than generally assumed, and to understand it we need a thorough macroeconomic analysis based on the book-entry nature of bank money and of the way it is issued and integrated with produced output.

4

From Monetarism to the New Classical Synthesis

The aim of this chapter is to show that both monetarism and the new classical synthesis fail to provide a satisfactory analysis of the working of our economies and of the way disorders may arise, disrupting the economic system taken as a whole.

The first section is devoted to a critical appraisal of monetarism, starting from Friedman's (1956) restatement of the quantity theory of money. In particular, we will show that the quantity equation on which this theory rests is tautological, and that the concept of 'quantity of money' is ill founded and completely at odds with the true nature of bank money. As a matter of fact, the whole conception of money advocated by monetarism contrasts with the reality of double-entry bookkeeping and the immateriality of bank money.

New classical economics fares no better. As shown in the second section, new classical analysis was developed in the 1970s in order to counteract the widespread use of Keynesian models. To this effect, it provides new general equilibrium models explaining business cycles consistently with the microeconomic approach typical of neoclassical analysis. The rational expectations hypothesis plays a crucial role in these models and is symptomatic of the attempt to identify the causes of economic crises in irregular external shocks and imperfect information.

The transition from real business cycle analysis to the new classical synthesis analysed in the third section marks the resolve to integrate monetary policy and market imperfections into a theoretical framework where exogenous technological shocks are seen as the main cause of economic fluctuations. As a result, imperfect markets and nominal rigidities are incorporated into real business cycle models in so far as they prevent an efficient response to real shocks, and monetary policy is accepted as a possible remedy when it reacts to this kind of shocks.

Whether in the form of real business cycle models or in that of monetarist or new classical models, the analysis proposed by neoclassical economists is based on the concepts of equilibrium and relative exchange, as well as on microeconomic foundations and on the dichotomy between real and monetary variables. Despite several attempts to incorporate monetary policy into various neoclassical models, money continues to have little or no impact and is still considered as a commodity or identified with a net financial asset. The true nature of bank money is foreign to neoclassical analysis in all its forms, which is why its models are but highly sophisticated constructs removed from the reality of our monetary economies.

Monetarism

In its modern version, monetarism has its origin in Friedman's 1956 restatement of the quantity theory of money and relies on two strong assumptions, namely that

- 1) Fisher's quantity equation can be transformed from a truism into a condition of equilibrium; and
- 2) money can be identified with a positive asset exogenously determined.

Let us consider these two assumptions in succession.

In the chapter of his 1956 book devoted to his restatement of the quantity theory of money as well as in his 1970 article entitled 'A Theoretical Framework for Monetary Analysis', Friedman attempts to give a theoretical foundation to his and Schwartz's empirical findings of 1963. He does so by going back to Fisher's (1911) transaction equation, which states that the quantity of money (M) times its velocity of circulation (V) is necessarily equal to the price level (P) times the number of transactions per unit of time (T): $MV = PT$ (Fisher 1911: 24). The following steps lead Friedman to introduce the income form of the quantity equation, $MV = Py$ (where P stands for the price index implicit in estimating national income at constant prices, y stands for national income at constant prices, and V stands for the average number of times the money stock is used in making income transactions), from which he derives the demand function for money, $M = kPy$ (with k representing the ratio of money stock to income).

Most of Friedman's effort in this connection is devoted to working out a model that reconciles the quantity theory of money with Keynes's

income–expenditure theory. As interesting as such an endeavour may be, we shall leave this question aside and concentrate our attention on the logical possibility to transform the quantity equation into an equilibrium condition setting the basis for a theoretical framework for monetary analysis. Prior to any attempt to work out a general model encompassing neoclassical and Keynesian approaches, it has to be established if Fisher's (1911) equation is appropriate in its role as the cornerstone of economic theory. Friedman (1970) apparently is convinced that this is indeed the case. As a matter of fact, he claims that, even though the quantity equation is nothing but a tautology, it nevertheless 'is a useful device for clarifying the variables stressed in the quantity theory' (Friedman 1970: 3).

A tautology is defined as a repetition of an idea, a statement or a word. However, a distinction should be made between tautologies that are useless repetitions and those that increase our knowledge. Truisms pertain to the first category, logical-analytical statements to the second. Thus, for example, the identity $a = a$ is a mere truism, a self-evident truth with no heuristic value whatsoever, while $(a + b)^2 = a^2 + 2ab + b^2$ is a meaningful tautology improving our understanding in the field of algebra. In the case of Fisher's equation it is certain, as also alleged by Friedman (1970), that each term is the repetitious substitution of the other. 'Indeed, it is a tautology, summarized in the famous quantity equation, that all changes in nominal income can be attributed to one or the other' (ibid.: 3). The first point in need of clarification is whether the quantity equation belongs to the category of meaningful tautologies or to that of truisms. Friedman himself leads us to the correct answer by observing that Fisher's equation 'like the other quantity equations I shall discuss, are intended to be *identities* – a special application of double-entry bookkeeping, with each transaction simultaneously recorded on both sides of the equation' (ibid.: 6, italics added). As the quantity equations result from the simple application of double-entry bookkeeping, the first term of every equation is just the repetition of the second, and vice versa. Needless to say, if we know the number of transactions carried out in a given period of time and their price (PT), we also know the amount of money spent in the same period of time (MV).

No causal relationship exists between the two terms of an identity, a conclusion that should have led Friedman to relinquish any hope of transforming the quantity pseudo-equation into a theory-analytic one. Instead, he remained faithful to his initial project and advocated the thesis that, even though it is essentially a tautological identity, the quantity equation (whatever form it takes) can be considered as a condition of

equilibrium, because each term may be evaluated independently of the other. 'However, as with the national income identities with which we are all familiar, when the two sides, or the separate elements on the two sides, are estimated from independent sources of data, many differences between the two sides emerge' (ibid.: 6). Unfortunately, Friedman's contention is hardly in keeping with the principles of logic. An identity is the strongest possible relationship between two terms defining the same entity, to wit, a relation that holds between any entity and itself or the quality of being united into one. The very idea that any difference whatsoever may exist between the two terms of an identity is a *contradictio in adjecto*, and must be rejected unconditionally.

Despite this uncompromising conclusion could we venture to rescue Friedman's thesis by claiming that although the monetarist quantity equations are identities in theory, in practice their terms can differ, since they may be derived from different sources? To answer yes to this question would mean to confine logic to the abstract world of theories (as if practice could ever exist without any supporting theory), as well as to deprive the status and logic of double-entry bookkeeping from all its rigour. As attractive as this prospective might appear to those allergic to logical strictness, such answer is entirely inadequate if not misleading. In fact, any numerical difference between the two terms of an identity either can be immediately eliminated or can become the mark of an incorrect interpretation of the identity itself. Fisher's quantity equation cannot be misunderstood, and Friedman correctly defines it as an identity. What he fails to recognize is that, once it is grasped that it is a self-evident truth, it can no longer be transmogrified into a condition of equilibrium.

Whether one likes it or not, our conclusion is straightforward: the quantity theory of money lacks logical rigour, and cannot provide a satisfactory explanation of the principles founding and governing our monetary systems. This is not tantamount to saying that monetary pathologies are not the source of economic crises. Money, credit, and capital are at the core of a theory of crises, and the banking system is no less important for their macroeconomic instituting and role. The failure of the quantity theory of money brings to the fore the need for a new approach to monetary theory. In particular, it is necessary to start afresh by considering the all-important but neglected problem of what money is and how it can become associated to produced output.

[A]ll the adherents of the monetary theory of the trade cycle have sought an explanation either exclusively or predominantly in the

superficial phenomena of changes in the value of money, while failing to pursue the far more profound and fundamental effects of the process by which money is introduced into the economic system, as distinct from its effect on prices in general.

(Hayek 1933/2008: 6)

This is the challenge that economists have to face if they aim at bearing out the nature of macroeconomic crises.

Let us now introduce few considerations as to the second assumption of monetarism.

Friedman and his followers assume that money has a positive intrinsic value, and that its stock or quantity is exogenously determined by the monetary system. 'The factors determining the nominal quantity of money available to be held depend critically on the monetary system' (Friedman 1970: 10). Consistently with their acceptance of the neoclassical framework of analysis, monetarists accept the homogeneity postulate and the money neutrality axiom. Yet they attempt to reconcile the neoclassical assumption that only real variables matter with the indisputable evidence that money may influence equilibrium. They do so by claiming that, whereas in the short run changes in the money supply affect both real and nominal output, in the long run changes in real output are independent of fluctuations in the money stock. As Friedman (1970: 27) noted, 'I regard the description of our position as "money is all that matters for changes in *nominal* income and for *short-run* changes in real income" as an exaggeration but one that gives the right flavour of our conclusions'.

In fact, neoclassical economists are trapped between their belief in general equilibrium analysis and the necessity to account for the peculiar role played by money. Were they to remain faithful to Walras's conception of general equilibrium, they would have to turn down any attempt to give money a role in the determination of real variables. Conceived as a veil or as a medium of exchange, money is not supposed to alter the relation of exchange established between goods and services by the interplay of their supply and demand on the commodity market, as well as any other relationship between real variables as determined on the other markets. Its presence may have a practical relevance, but it has no one from a purely theoretical point of view, and it is difficult to see how it could modify, albeit momentarily, the equilibrium resulting from the simultaneous solution of the equations defined by general equilibrium analysis. On the other hand, as observed by advocates of the quantity theory of money, it is hard to maintain that

changes in the money supply have no impact whatsoever on economic agents' behaviour. The dilemma cannot be avoided by switching from an alternative to the other according to what best suits the researcher confronted with a given problem. Logical coherence has to be preserved, which would not be the case if money were defined as a superfluous device (a veil) capable to disrupt equilibrium. Either money is inessential, and real variables are all that matters, as in Walras's and Debreu's general equilibrium models, or money is an essential element of our economies, and cannot be set aside at will. Monetarism is an attempt to reconcile these two reciprocally exclusive paradigms. No wonder that it has proven incapable to provide a satisfactory explanation of economic and financial crises.

The most powerful argument against monetarism rests on the nature of bank money. The assumption that money is a positive asset issued by banks is at odds with factual observation, and overtly inconsistent with the principle of double-entry bookkeeping. If money were a stock of physical objects, say a quantity of marbles, it would indeed circulate from the payer, A, to the payee, C, in a unidirectional flow opposite to the flow of goods and services purchased by A. Payments being carried out by banks on behalf of their clients, A's bank, B, would thus debit A and credit C. In order to avoid useless complications, let us assume that C, the seller, and A, the purchaser, are clients of the same bank. The payment of C in money-marbles would be entered in B's ledger as shown by Table 4.1. According to this book-entry representation, the x units of money-marbles paid by A are credited to C, who is the new owner of the deposit of marbles formed in B. The payer is debited and the payee is credited with x marbles, which are at the same time the medium and the object of the payment.

The frustrating aspect of this interpretation is that it does not apply to the real world of monetary economies, where bank money is immaterial. Marbles are physical objects and, as such, they may be transferred from one economic agent to another. They may also be deposited with bank B, and the transfer from A to C reduced to the transfer of the ownership

Table 4.1 The use of money-marble as unit of payment

Bank B			
Assets		Liabilities	
Purchaser A	− x units of money-marble	Seller C	+ x units of money-marble

over the deposit. Yet, bank money is no physical object, and its emission must comply with the principle of double-entry bookkeeping. Correctly understood, this principle states that a bank cannot credit a client without debiting her/him and, reciprocally, it cannot debit a client without crediting her/him simultaneously. Each single economic agent must at the same time be credited-debited or debited-credited for the same amount of money units; this is the deep significance of that principle. Each entry must be double, which is entirely different from saying that each entry must be balanced by another entry. Double-entry bookkeeping does not mean that the credit of a given economic agent (C) must be matched by the equivalent debit of another economic agent (A), but that both economic agents, A and C, must be debited and credited or credited and debited. Double-entry is indeed double and not simple, as would be the case if money was a material object.

All the concepts proper to monetarism indicate that money is identified with a physical asset. Mass, stock, or quantity refer to the physical world, and so does the concept of velocity and that of circulation. Even though today central banks are getting sceptical about the traditional way of accounting for the quantity of money and for its impact on monetary disequilibria, no alternative has been worked out let alone adopted so far. Money is still widely defined as an asset whose transfer from the payer to the payee allows for the settlement of a transaction that, substantially, is nothing else but a relative exchange between the goods and services sold by the payee and the money given up as counterpart by the payer. Now, what the advocates of monetarism will never be able to explain is how a bank could issue such an asset.

Nobody can dispute the fact that money is bank money, which means that the first step in need of explanation is its emission by banks. Logically, the analysis must start from *tabula rasa*, its aim being that of explaining how money first appears on the scene without calling for an act of faith. Unless we switch from science to metaphysics, we have to accept the fact that nothing comes out of nothing, and if the capacity for a bank (as well as for any other human institution) to create a positive asset is excluded, the analysis must concentrate on the possibilities opened up by the discovery of double-entry bookkeeping. Its substantial achievement is to allow banks to issue money as an asset-liability. The idea of money creation is acceptable only if it is coupled with that of the instantaneous destruction of money, a requirement that double-entry bookkeeping alone satisfies.

By debiting and crediting A with x money units and crediting-debiting C at the same time and for the same amount, bank B creates and destroys

x units of money, as noted in Chapter 1 already. Contrary to what is usually believed, therefore, money is not an asset or a commodity, has no intrinsic value, cannot be held or hoarded, does not circulate among agents more or less rapidly, and does not flow in the opposite direction of the goods and services purchased by the payer. The fact that today the great majority of economists still reason as if money were a net asset clearly shows how far they are from a correct understanding of the principle of double-entry bookkeeping. Why is it so? Why is it so difficult to grasp the innermost nature of money?

The answer is far from being straightforward. Smith's distinction between money and money's worth or between nominal and real money, Marx's conception of money as a form of value, Walras's concept of *numéraire*, and Keynes's wage units are far reaching intuitions that have not been followed up. Was it because they are too abstract? Very likely not. If abstraction were a criterion to choose between different analyses, quantum mechanics would have never become the leading theory in the field of elementary particles, and this would have been true for many other theories in numerous research fields. Was it because it seems absurd to claim that money has no intrinsic value whatsoever while it can nevertheless be used as a means of payment? The answer should be no again, as it is not too difficult to understand that money does not need to be the 'object' of a payment in order to carry it out or 'vehiculate' it. Arguably, three elements may have played a determinant role in diverting economists from a clear perception of what money really is: the identification of money with what is used to represent it, general equilibrium analysis, and the distinction between stocks and flows.

From pebbles to silver and gold, money has for ages been identified with a more or less valuable material it was supposed to be made of or convertible in. When double-entry bookkeeping was discovered and for centuries after its adoption, bank money was defined in terms of gold or silver, and it is only in very recent times that central banks have severed any link between bank money and gold reserves. The propensity to make money coincide with a material object, preferably a valuable one, is anchored in the history of humanity, and it is not surprising, therefore, that the new insight provided by double-entry bookkeeping has not yet led us to discard any material conception of money. However, the passage from gold to bank money together with the increasing dematerialization of book-entries (an electric impulse being enough to enter billions in a bank's ledger) should arise doubts as to the correctness to identify money with any of its physical representations. This is

to say that analysis must go deeper than appearances in its search for a satisfactory definition of money, and double-entry bookkeeping provides all that it is needed to do so, on condition that it is not reduced to a concept of equilibrium between two entries relating to two different subjects.

It is at this stage that general equilibrium analysis exerts its negative influence, since by advocating the idea that exchanges can only be relative it has been able to impose the concept of equilibrium to almost the entire economics profession. Interpreted according to this concept, double-entry bookkeeping becomes a principle establishing the necessary equilibrium between two distinct entries, and not the necessity to enter twice each of them. The choice is between a simple double-entry, where the debit of A is matched by the credit of C, and double double-entry, where both A and C are submitted to the principle of the necessary equality between assets and liabilities. In his *Elements of Pure Economics*, Walras (1874/1984: 230) writes: 'This is double entry bookkeeping. Its cardinal principle is *never to enter a sum to the debit or credit of an account without simultaneously entering the amount to the credit or debit of some other account.*'

Walras's conception of double-entry bookkeeping is too linear and cannot account for the emission of bank money. The balancing between the debit of A and the credit of C is the balancing between two different entries each one 'simple', and applies only to the transfer of an asset. In the case of commodity-money, say gold, bank B would carry out the payment of its client A by debiting her/him and crediting C for a given amount of gold. B would transfer part of the ownership over its gold deposits from A to C, so that C would become the new owner of the gold deposit previously owned by A. Yet, since money is not a commodity at all, B cannot issue it as an asset, but has to rely on double-entry bookkeeping in order to issue the money required to carry out the payment of C on behalf of A. In other words, since B cannot create the object of A's payment, it has to confine itself to the emission of a numerical means of payment. As created by banks, money is but an empty vehicle, circulating instantaneously from the payer to the payee and back to the bank issuing it (see Chapter 1).

Walras's great intuition is that of conceiving money as a *numéraire*, which should have led him to the definition of money as a purely numerical form. Unfortunately, the French economist was not true to his intuition, and identified the *numéraire* with a commodity. 'The commodity in terms of which the prices of all the others are expressed is the "*numéraire*"' (ibid.: 161). Once money is thus conceived, its exchange

with produced output can only pertain to the category of relative exchanges. The payment in gold of the goods and services purchased by A is the relative exchange between two different commodities, and can be represented using double-entry bookkeeping in its weakest sense. The use of this 'simple' version of double-entry bookkeeping, however, is not possible when money is just a numerical means of payment. If, when issuing money, bank B were to credit A without immediately debiting her/him, the result would be the creation of a net asset to A's benefit. Since banks cannot create income, A must simultaneously be credited and debited with the amount of money issued by B. The same is true for C, who is indeed paid by means of the money units issued by B, but whose payment has not a simple number as its object or 'content'. What C finally gets in exchange for her/his goods or services is not a sum of money units, but a positive income, that is, a part of current output. C is credited-debited in money units, and becomes the owner of a positive bank deposit formed in B.

It is production that provides for the 'content' of money. The role of banks is to create the numerical vehicle conveying current and future outputs, and they do so using double-entry bookkeeping in its strong sense. Besides the two reasons mentioned so far, what makes it difficult to become aware of this strong sense is the fact that bookkeeping entries into banks' ledgers relate to stocks, while money is a flow. From the point of view of stocks (income, capital), the use of double-entry bookkeeping in its weak sense is apparently enough to account for the result of payments. When A pays C, s/he does so by asking her/his bank to debit her/his income account and credit that of C. If we consider only the banking aspect of the payment and we reason in income terms only, it seems that the whole transaction amounts simply to the debit of A and the credit of C.

In appearance, it is only if we aim at representing the intervention of money that we have to recur to double-entry bookkeeping in its strong sense. Being an instantaneous flow, money conveys A's payment to C, who is thus credited by bank B. The object of C's credit is a positive amount of income deposited with B, and defines an equivalent amount of current output. The instantaneous flow of money from B to A and C, and back to B is what allows for A being debited and C credited with that sum of income. This means that, in money terms, A and C are credited-debited. If this was not so, to wit, if A was simply debited and C credited in money terms, A would lose a sum of money units gained by C, which is clear nonsense as bank money is valueless and immaterial, and cannot be the term of a relative exchange.

Let us show in some details what happens when A asks her/his bank to pay C on her/his behalf for her/his purchase of a given amount, say x units in value terms, of produced goods and services. In order for this payment to be carried out, A must dispose of an income, a condition that is fulfilled either if A owns a claim on B's deposits or if s/he benefits from a loan. Let us consider the first alternative, A being the owner of a bank deposit of x units obtained through her/his labour services. When A asks B to pay C, s/he is credited-debited of x money units, which is perfectly justified by the fact that A gives up her/his claims on a bank deposit and purchases C's goods and services. A is thus credited with x money units for her/his sale on the financial market, and debited of the same amount of money units for her/his purchase on the commodity market. As far as C is concerned, her/his credit-debit in money terms is explained by the sale of goods and services and the purchase of claims on bank deposits. As soon as C is credited with x money units for her/his sale on the commodity market, s/he is debited of the same sum and becomes the owner of an equivalent bank deposit. A's and C's credits-debits occur simultaneously. Through double-entry bookkeeping a sum of money is thereby instantaneously created and destroyed on A and on C, this circular flow of money accounting for the cancellation of A's bank deposit and the constitution of a new bank deposit owned by C as well as for the transfer of the economic ownership over a fraction of current output from C to A.

What is crucial to understand is that the debiting of A and the crediting of C occur through the intermediation of bank money, and this intermediation implies the use of double-entry bookkeeping necessarily. Economic transactions are always conveyed by money, and money can but be submitted to this principle. Now, as our example clearly shows, double-entry bookkeeping applies both to the monetary and the real aspects of economic transactions. Indeed, A and C are credited-debited and debited-credited also in real terms, which should not come as a surprise to the reader, because it has already been shown (see Chapter 2) that each agent's sales (purchases) are always necessarily balanced by her/his simultaneous purchases (sales). Unfortunately, by keeping separate the monetary from the real results of payments, economists have missed the chance to understand that economics rests on the identity (established by production) between the monetary and the real worlds. The temporary triumph of the neoclassical dichotomous conception of these two worlds has seriously hampered the development of monetary economics and prevented the correct understanding of double-entry bookkeeping, even though it may be argued that general

equilibrium analysis is not inconsistent with this fundamental principle of accounting. Let us show it.

The neoclassical dichotomy is particularly in line with the definition of money as a veil. According to the homogeneity postulate, first sketched by Hume (1826) and then taken over by Fisher (1911) and the advocates of the quantity theory of money, '*one of the normal effects of an increase in the quantity of money is an exactly proportional increase in the general level of prices*' (Fisher 1911: 157). This amounts to saying that only relative prices matter, monetary variations affecting their nominal expression only. Being a veil, money plays no essential role: it does not alter the relations established through direct exchange, and is never demanded for its own sake. As a consequence, money remains essentially extraneous to the world of goods and services. Money is considered neither as the numerical form of produced output, nor as a valuable asset, neither as a numerical vehicle, nor as the counterpart of output. On these conditions, payments are basically accounted for in real terms only, and double-entry bookkeeping is reduced to its simplest form, where the purchaser is debited and the seller is credited.

In fact, a doubt arises in this regard, because according to general equilibrium analysis direct exchanges occur simultaneously: if this were indeed the case, each agent would finance her/his purchases through an immediate sale. Money would not interfere with direct exchange, and would itself be offered and demanded simultaneously and for the same amount. We thus get the impression that by defining money as a veil and by advocating the simultaneous determination of equilibrium on every market, neoclassical analysis is perfectly consistent with the correct definition and the use of double-entry bookkeeping. In reality, this conclusion is too hasty, as in the analytical framework proposed by general equilibrium advocates there is no room for the formation of capital. The lack of integration between money and goods is fatal, because if money is merely a veil, no bridge can be built between present and future. The implementation of double accounting within neoclassical analysis ends up with the simultaneous debit and credit of every economic agent on the commodity market, a constraint far too strong to be consistent with the working of our capitalist economies.

The same result is reached when money is identified with a net asset. If money is said to enter general equilibrium as a valuable asset, monetary payments become but a particular case of relative exchange. According to this definition, money would not substantially differ from any other good, and in its exchange with output it would 'move' along a one-way path, from the purchaser (A) to the seller (C). Yet, even in this

case it is possible to claim that payments take place respecting the principle of double-entry bookkeeping. This is so because general equilibrium rests on the simultaneous clearing of all markets, which implies that every agent implicated in a relative exchange is at the same time a seller and a purchaser, whether money intervenes as a medium of exchange or not. As in the previous case, however, the consistency of neoclassical analysis with the principle of double accounting is fruitless, because the implementation of this principle in that framework leaves capital totally unexplained.

Eventually, neither the quantity theory of money nor general equilibrium analysis provide an adequate explanation of our economic systems, since none of them is able to reconcile the requirements of double-entry bookkeeping with the essence of a system based on the generalized use of bank money.

New classical analysis and the rational expectations hypothesis

New classical economics developed in the 1970s as a reaction against the generalized use of Keynesian econometric models whose predictions proved incorrect in light of the simultaneous presence of high inflation rates and growing unemployment in the US economy. The US depression in the 1970s and the failure to predict it were seen as a good opportunity to discredit Keynesian economics and blame it for failing to incorporate general equilibrium analysis and understand the importance of microeconomic foundations. According to new classical economists, macroeconomic models have to be based on microeconomic foundations, namely on behavioural relationships accounting for consumers' and firms' maximization, and they must rest on the principle of market clearing and relative prices determination. They accused their Keynesian colleagues to have failed to properly incorporate expectations in their models, and to have discarded dynamic, intertemporal general equilibrium analysis. 'The failure of existing [Keynesian] models to derive restrictions from any first principles grounded in economic theory is a symptom of a deeper and more general failure to derive behavioral relationships from any consistently posed dynamic optimization problems' (Lucas and Sargent 1978: 75).

The aim of new classical economics was to provide new general equilibrium models replacing those advocated by the Keynesian, and capable to explain business cycle fluctuations in accordance with the main tenets of neoclassical analysis. As Lucas and Sargent (1978: 78)

claimed, the research line pursued by the supporters of new classical economics involved 'the attempt to discover a particular, econometrically testable equilibrium theory of the business cycle, one that can serve as the foundation for quantitative analysis of macroeconomic policy'.

Given a theoretical framework in which markets clear and economic agents optimize, it is assumed that information is limited and unevenly distributed. The rational expectations hypothesis is then introduced in order to allow for objective probability distributions to be determined consistently with what general equilibrium analysis requires. As a consequence, business cycle fluctuations are explained by referring to how economic agents react to unforeseen changes in the price level owing to disturbances in aggregate demand. 'The central idea of the equilibrium explanations of business cycles sketched above is that economic fluctuations arise as agents react to unanticipated changes in variables which impinge on their decisions' (ibid.: 80). To put it shortly, by assuming that markets are perfectly competitive, new classical economists claim that fluctuations in business cycles should be attributed to imperfect information rather than to the failure of wages and prices to immediately adjust to monetary shocks.

Well aware that money is far from being a simple veil and that monetary disturbances may seriously disrupt our economies, new classical economists elaborated a new series of models based on the rational expectations hypothesis advocated by Lucas (1972, 1977). As maintained by Goodfriend and King (1997: 241), derived from Muth 1961 suggestion that individuals form expectations optimally, this hypothesis 'is a natural extension of the neoclassical principle that the economy is inhabited by rational, maximizing agents'. Indeed, some sort of rational expectations hypothesis is also present, more or less explicitly, in Friedman's version of the quantity theory of money: by emphasizing the relevance 'of a *verifiable commitment* by the central bank to a non-inflationary policy' (Woodford 2007: 4), monetarists were stressing 'the role that commitment to a policy rule could play in creating the kind of expectations needed for macroeconomic stability' (ibid.: 4). Analysing the role of monetary policy in his 1968 article published by the *American Economic Review*, Friedman claims that wages and prices respond smoothly to changes in the money supply if expectations are sound, and suggests that in this case sustained inflation is neutral in the long run.

In short, the rational expectations hypothesis is the assumption that economic agents react rationally to changes, provided that these are regular and information reliable.

Insofar as business cycles can be viewed as repeated instances of essentially similar events, it will be reasonable to treat agents as reacting to cyclical changes as 'risk', or to assume their expectations are *rational*, that they have fairly stable arrangements for collecting and processing information, and that they utilize this information in forecasting the future in a stable way, free of systematic and easily correctable biases.

(Lucas 1977: 15)

More specifically, economic agents are assumed to be influenced by expectations about future changes consistently with the principles established by economic analysis. '[Lucas, Sargent and their co-authors] proposed that expectations be modelled, not by *any* specified function of past experience (as in earlier monetarist efforts), but instead by assuming that people's expectations coincide at all times with what one's economic model implies should occur' (Woodford 1999: 19). Needless to say, the economic model economic agents are supposed to use in formalizing their expectations is the one advocated by Lucas and his followers. New classical economics is thus supposed to faithfully represent the principles governing the theoretical framework as well as the practical working of our economies.

The rational expectations hypothesis itself rests on two unlikely assumptions, namely that, despite their subjective and objective differences, economic agents work out their expectations in the same way, and that they do so by referring, mostly implicitly, to the principles settled by mainstream economics. We are thus asked to believe at the same time that behavioural economics lays down the foundations of both micro- and macroeconomic analysis, and that economic agents' decisions are determined by a set of expectations they do not master. In other words, we are told that expectations are imposed to the very economic agents whose behaviour is supposed to determine the principles governing our economies.

It should be superfluous to stress how artificial it is to assume that economic agents confronted with expected or unexpected changes react in a similar way. Human behaviour is highly unpredictable and depends on too many variables to endorse such an assumption. Goodhart (2007: 4) recognizes it, albeit only in the case of very unsettled circumstances: 'during turbulent periods, whether of severe deflation or inflation, expectations will not be anchored, will differ quite markedly from person to person, and be subject to potentially rapid and sharp revision'.

If so, then a question arises naturally: why is the rational expectations hypothesis given such a prominence within neoclassical analysis? The answer is straightforward: without this hypothesis, neoclassical modelling would result almost impossible. If no strong assumption is introduced in order to drastically reduce the number of variables, the use of models to reproduce or mimic reality would be doomed to failure. The rational expectations hypothesis plays this role to the great advantage of neoclassical analysis, as not only it seems to allow for the use of neoclassical (or new classical) models, but it also imposes the theoretical choice of Lucas and his followers as the standard required in order to establish whether or not expectations are rational. As a matter of fact, the rational expectations hypothesis has mainly been exploited by general equilibrium analysis, from its new classical to its real business cycle version. 'Although rational expectations were introduced into macroeconomics to study the links between real and nominal variables, its implications were more systematically worked out within the real-business-cycle research program' (Goodfriend and King 1997: 242). The acceptance of the rational expectations hypothesis, however, is not limited to this stream of economic analysis, as it can easily accommodate whatever theory is chosen as a reference. Muth's hypothesis has gained so wide an acceptance that, as claimed by Woodford (2008), 'it is now routine both in positive interpretations of macroeconomic data and in normative analyses of possible economic policies to assume rational expectations on the part of economic decisionmakers, in accordance with the methodology introduced by the New Classical literature of the 1970s' (Woodford 2008: 8–9).

From real business cycle analysis to the new classical synthesis

While new classical economics may be deemed as a neoclassical attempt to incorporate nominal rigidities, imperfect information, and monetary disturbances into general equilibrium analysis, real business cycle economics refocuses attention on the main tenets of Walras's and Debreu's economic models. In the words of Plosser, a founding father of real business cycle analysis:

Real business cycle models view aggregate economic variables as the outcomes of the decisions made by many individual agents acting to maximize their utility subject to production possibilities and resource constraints. As such, the models have an explicit and firm foundation

in microeconomics. More explicitly, real business cycle models ask the question: How do rational maximizing individuals respond over time to changes in the economic environment and what implications do those responses have for the equilibrium outcomes of aggregate variables?

(Plosser 1989: 53)

Monetarists' considerations about the role played by money and monetary policy, the Phillips curve, wage and price rigidities, as well as market inefficiencies are no longer important factors, and a series of models is proposed, where rational economic agents operating on essentially efficient markets maximize their utility and their profit. Emphasis is entirely on an intertemporal optimization approach to consumption and investment derived from the utility-maximizing and the profit-maximizing decisions of consumers and firms within a general equilibrium framework. As Blanchard (2009: 211) puts it, '[t]hree principles guided the research: explicit micro foundations, defined as utility and profit maximization; general equilibrium; and the exploration of how far one could go with no or few imperfections'.

As a matter of fact, as in King and Plosser (1984), real business cycle models have also been used to investigate the correlation between money and output, but only to the extent that it is influenced by real productivity shocks. In contrast with previous speculations about the origin of economic fluctuations, real business cycle economists do not believe that monetary shocks and market inefficiencies can lead to relevant and permanent variations of the business cycle both in the short and in the long run. They maintain instead that exogenous technological shocks are the main cause of fluctuations, and put them at the core of dynamic stochastic general equilibrium models accounting for permanent changes in productivity.

[R]eal per capita output, as well as many other economic time series, behave as if they have random walk components (much like the log of stock prices). Random walks have the important property that there is no tendency for the process to return to any particular level or trend line once displaced. Thus, unpredicted shocks to productivity permanently alter the level of productivity.

(Plosser 1989: 59)

Real business cycle analysis has been criticized on the ground that, by resting on the neoclassical dichotomy of the nineteenth century,

it leaves no room for monetary policy and market imperfections to play any essential role in affecting business fluctuations. For example, Blanchard (2000: 18) recognizes that the ‘methodological contributions of the Real Business Cycle approach, namely the development of stochastic dynamic general equilibrium models, have proven important and have been widely adopted’, but emphasizes the fact that ‘the initial propositions that money did not matter, that technological shocks could explain fluctuations, and that imperfections were not needed to explain fluctuations, have not held up: the empirical evidence continues to strongly support the notion that monetary policy affects output’ (ibid.: 18).

Yet, extreme positions have softened lately, and even the most faithful adherents to the real business cycle approach do not reject aprioristically the idea that ‘an economy in which monetary fluctuations induce real inefficiencies – indeed an economy in which money has value – must be one that operates under missing markets and nominal rigidities that make changes in money into something other than mere units changes’ (Lucas 2003: 3). Indeed, it is widely accepted that if imperfect markets and nominal rigidities are such as to ‘prevent the economy from responding efficiently to real shocks [they make it possible] that a monetary policy that reacts to real shocks in some way can improve efficiency’ (ibid.: 3). Be that as it may, it is clear that, in any of its versions, real business cycle analysis is built on general equilibrium principles, and is substantially in line with Debreu’s axiomatic theory of value (relative prices). The critical observations that we raised against Walras’s general equilibrium model and against the new classical approach can be applied to dynamic stochastic general equilibrium models developed by real business cycle theoreticians. Instead of dwelling again on these criticisms, let us explain what distinguishes the neoclassical from the quantum theoretical concept of money neutrality.

The reader might be tempted, in fact, to claim that there is no substantial difference between maintaining that money is instantaneously created and destroyed in each payment, and that money is but a veil in a world where goods and services exchange among themselves. As we pointed out, according to quantum macroeconomics money is a flow conveying or ‘vehiculating’ goods, services, and financial claims from the payee, C, to the payer, A, and vice versa. C sells goods and services and purchases a claim on a bank deposit, while A purchases C’s goods and services and disposes of equivalent bank deposits. Money enters the exchange only as an immaterial means of payment, and leaves immediately its place to what has been exchanged thanks to its

intermediation. In general equilibrium analysis, money is also a simple means or medium of exchange, something that mediates direct exchange between goods and services, but this time money is considered as a veil that does not alter the rate of exchange between goods and services as determined by the adjustment of their supply and demand.

The difference between the two approaches is determined by the fact that according to neoclassical analysis money is not a specific element of capitalist economies (any particular good may be chosen, if necessary, to play the role of money without any impact on relative prices), whereas in quantum macroeconomics money is central, no economic system being conceivable in its absence. It is true that, as issued by banks, money disappears at the very moment it is created. Yet, an instant is enough to give a numerical form to produced output and to generate a positive amount of income. If physical goods and services were not given a numerical form, they would remain totally heterogeneous. Physical heterogeneity is the first obstacle economic analysis has to overcome. Neoclassical analysis fails to satisfactorily address this problem for lack of a numerical standard. What the advocates of general equilibrium fail to see is that bank money provides the required numerical standard and that, through double-entry bookkeeping, an absolute exchange can take place, which transforms current output into a sum of income. Even if it remains true that, once payments have been carried out, money disappears, its instantaneous presence is enough to make physical goods and services commensurable, to transform them into income, and to build a bridge between present and future, that is, to account for the formation of capital. None of these tasks can be fulfilled starting from direct exchange (see Schmitt 1984a, 1996).

To be sure, the concept of money neutrality has two different meanings according to whether it is referred to general equilibrium or to quantum theoretical analysis. In a general equilibrium framework, money is essentially neutral because real variables alone matter, and relative prices are not subject to money variations. The neutrality of money is just a corollary of the centrality of direct exchange, just an alternative way of stating the substantial irrelevance of money. If, irrespective of what is implied by the theory, monetary disturbances have repercussions on the real world, order has to be restored by reducing the importance of money. According to the quantum analysis advocated in this volume, by contrast, money remains central in all circumstances. Its neutrality does not mean that it is essentially superfluous, but that its use must comply or be consistent with the principle of double-entry bookkeeping as applied to money, income, and capital. Anomalies such

as inflation and deflation derive from the lack of conformity between the present structure of monetary payments and the essential difference existing between money, income, and capital. To restore order means to reform the system of payments so that double accounting respects this triple distinction, thereby preventing those anomalies characterizing the lack of money's neutrality.

Starting in the 1990s, an increasing convergence between neoclassical and Keynesian traditions has been taking place, which led to what Goodfriend and King (1997) call the new neoclassical synthesis. Basically, the advocates of this new synthesis take over some of the main tenets of monetarism, new classical economics, real business cycle analysis, Keynesian and New Keynesian analyses, and work out a series of models in which these views are all given a role to play, in an attempt to overcome their differences by incorporating them in the same mould. Mostly, these models are dynamic stochastic general equilibrium models derived from new classical and real business cycle analyses, where rational expectations and intertemporal optimization occupy a relevant place. The Keynesian and New Keynesian contributions to these new models are essentially the incorporation of wages and prices stickiness, price-setting firms being modelled as monopolistic competitors.

The New Neoclassical Synthesis is defined by two central elements. Building on new classical macroeconomics and RBC [real business cycle] analysis, it incorporates intertemporal optimization and rational expectations into dynamic macroeconomic models. Building on New Keynesian economics, it incorporates imperfect competition and costly price adjustment. Like the RBC program, it seeks to develop quantitative models of economic fluctuations.

(Goodfriend and King 1997: 255)

The new neoclassical synthesis aims at overcoming the methodological differences between micro- and macroeconomics by modelling Keynesian key features using the principles of general equilibrium analysis, which is made to switch from its traditional static general equilibrium to its intertemporal general equilibrium version. As stressed by Woodford (1999: 29), in practice 'this means that the methodology of the new synthesis is largely that of the real business cycle literature, even though wage and price rigidities are allowed for, and the determinants of (individually) optimal wage and price-setting decisions are modeled in detail'.

The generalized belief that macroeconomic analysis has to be based on microeconomic foundations is confirmed here by the fact that intertemporal general equilibrium models are supposed to reproduce both short- and long-run responses to economic disturbances. Adoption of the rational expectations hypothesis by Keynesians and New Keynesians is also a mark of the prevalent influence of general equilibrium analysis, in both its new classical and real business cycle interpretations, within the new neoclassical synthesis. Yet, the use of this hypothesis does not imply that stabilization policies are necessarily considered as ineffective. The introduction of wage and price rigidities into dynamic stochastic general equilibrium models has important consequences for the way wage and price-setting decisions are taken, and for the impact of stabilization policies.

According to those supporting the new neoclassical synthesis, economic disturbances leading to business cycle fluctuations are essentially of a technological kind, as claimed by real business cycle analysis. However, they might also include disturbances generated by government policies, by changes in preferences, or by a variety of other factors affecting aggregate demand. Even though the monetarist claim that business cycle fluctuations can largely be attributed to exogenous variations in the money supply is widely disputed, monetary policy is far from being considered as irrelevant, for the equilibrium reaction to real disturbances may depend substantially on these policies. Furthermore, if we observe that central banks are largely believed to have the capacity to control inflation, we might be tempted to conclude that money has still an important role to play. This is the case only insofar as the adherents to the new neoclassical synthesis accept the idea that the general level of prices is ultimately determined by monetary policy. The recognition of the role played by monetary policy, however, is not that of the relevance of money, which in some cases – as in Smets and Wouters (2003) – is not even taken into consideration.

Despite the large convergence on method, important disagreements remain as to the choice of the models best suited for quantitative policy advice. While mainstream economists broadly agree on the theoretical approach proposed by the new neoclassical synthesis, they are divided with regard to its practical implementation: ‘there is as yet little certainty about how best to specify an empirically adequate model of aggregate fluctuations’ (Woodford 2008: 14). The fact is that models are nothing other than poor simulations of a complex reality. As such, they are bound to provide only hypothetical constructions whose predictive power is highly uncertain and provisional. Confronted with this state of

affairs, mainstream economists either claim that more theoretical effort is required in order to further clarify the principles of macroeconomic theory, or concentrate on practical issues in an attempt to solve the problems undermining our economies through technical engineering (see Mankiw 2006). The long-lasting debate between economists advocating general equilibrium analysis and those supporting a Keynesian approach is not definitively set yet and, as recognized by some of the most authoritative members of both 'schools', there still are important theoretical and empirical issues in macroeconomics that have not been resolved. Woodford (2008: 19) is therefore right in claiming that '[o]ne can still hope for much more progress, and competition among contending approaches and hypotheses will almost inevitably be part of the process through which such progress can occur'.

The new neoclassical synthesis is the best representative collection of principles clearly at odds with the quantum theoretical analysis of monetary macroeconomics we advocate in this volume. To be sure, the choice between identities and equilibrium, absolute and relative exchange, integration and dichotomy, macro- and microfoundations, analytical and axiomatic theory, structural and behavioural economics, is not a matter of subjective preference. What is at stake is the possibility to understand and explain the very nature of our economic systems. As economic reality cannot simply be observed, but must be interpreted, we need an objective, analytical instrument guiding us in the search of the principles that will allow us to reach this understanding. As argued throughout this book, our economic systems are monetary, and bank money is their key feature. It thus follows that double-entry bookkeeping is the most appropriate analytical instrument economists have at their disposal to determine the specific laws of economics. If economic analysis is developed along the path traced by double accounting, it appears that economics can eventually be construed as a science of its own, and not, as done up to now, as a branch of physics and mathematics. Once money is conceived as a mere numerical form, as suggested by Walras's choice of the word *numéraire*, and as implied by the principle of double-entry bookkeeping, the task of economic analysis is settled: to show how physical heterogeneous products can enter into a rigorous relationship with this numerical form. Say's law and Keynes's identities result from the association between money and current output, and represent the fundamental principle on which macroeconomics rests. The idea that products may derive their numerical expression from direct exchange has its origin in the belief that individual preferences can be quantified, and an equilibrium found thanks to the appropriate

mathematical tools. In reality, preferences are economically meaningful only when referred to goods and services whose price (or value) has already been determined. The use of money as a numerical standard comes first, and the very object of economic analysis – the product – results from an absolute exchange that is made possible by the implementation of double-entry bookkeeping. General equilibrium analysis, in whatever of its forms, is an attempt to do away with money and to replace absolute with relative exchange. Its failure is already implicit in the decision to consider money as a secondary feature of economic theory, and to derive macro from microeconomics.

5

From Keynes to Post-Keynesian Economics

Keynes is unanimously considered as the ‘father’ of modern macroeconomics, and his followers have endeavoured to emphasize the crucial role played by macroeconomic concepts in seeking a theory capable of explaining both the orderly working of our economies and the insurgence of pathological states leading to the burst of economic and financial crises. Their interpretation of Keynes’s analysis, however, is not univocal, and some important differences exist between their approaches and the models they advocate. A classification of their contributions in the different schools of thought that call themselves Keynesian is not always easy and is meaningful only to the extent that it helps one to better understand the alternatives offered by each school. In this chapter, we will analyse the contributions of Keynesian, New Keynesian, and post-Keynesian economics in order to verify if they succeed in reaching a better understanding of the origin of crises than their neoclassical counterpart. In particular, we will show that, despite their emphasis on the role played by monetary disturbances and market imperfections, Keynesian economists of all schools fail to reach this goal.

Both Old and New Keynesians essentially fail to oppose the neoclassical interpretation of Keynes’s *General Theory*. Epitomized by their acceptance of Hicks’s (1937/1982) IS–LM interpretation, this failure has transformed Keynes’s analysis into a variant of general equilibrium analysis and has thereby deprived it of its revolutionary originality. Keynesian and New Keynesian economists have concentrated their efforts on finding adequate microeconomic foundations for their models and abandoned any attempt to look for the macroeconomic foundations of macroeconomic analysis. As our critique of Hicks’s IS–LM shows, the neoclassical interpretation of Keynes’s main contribution is

at odds with the logical identities on which the author of *The General Theory* founded his macroeconomic analysis.

New Keynesian economists follow the same neoclassical approach introduced into Keynes's analysis by Keynesian economists. Their contributions are essentially in line with those proposed by advocates of general equilibrium analysis, their difference being merely confined to the choice of hypotheses introduced in their models and to the relevance attributed to monetary shocks, market imperfections and asymmetric information. On the whole, despite their rejection of the market-clearing axiom and the representative agent introduced by neoclassical economists, New Keynesians accept that macroeconomics has to be grounded on microeconomic principles and that general equilibrium models should be used to mimic macroeconomic behaviour. By doing so, they give up any hope of escaping macroeconomics turning into a branch of microeconomics.

Are post-Keynesian economists more successful than their fellow economists in their attempt to build a true alternative to the general equilibrium framework? Only very partially. If one considers the emphasis some post-Keynesian authors put on the role of money and banks in a monetary production economy, one could be led to conclude that their approach is much closer to the message conveyed by Keynes's analysis than that of Keynesian and New Keynesian economists. However, post-Keynesian economics encompasses various and often divergent lines of thought. Even the definition of money differs from one strand of post-Keynesian thought to another. Thus, money is defined either as a means of payment or as liquid store of value; its value is explained by 'general acceptance' or is attributed to the intervention of the State, and its circulation is identified with income circulation or with the flow of a stock. Post-Keynesian economists seem to have completely lost sight of the conceptual and all-important distinction between money and credit. Besides their attempt to explain macroeconomics starting from behavioural considerations, their confusion between money and income is one of the main causes of their inability to provide a satisfactory alternative to the analyses advocated by their more orthodox fellow economists, self-designed heirs of Keynes.

Keynesian economics

The interpretation of Keynes's thought has moved along two different paths, one – dubbed as fundamentalist Keynesianism by Coddington (1976) – which has led to the development of post-Keynesianism, and

another that has given rise first to what we might label 'neoclassical Keynesianism' and then to New Keynesian economics.

Among fundamentalist Keynesians, Robinson, Kahn, and Shackle are the leading scholars. Their concern to keep out all general-equilibrium considerations as well as neoclassical-inflected interpretations of Keynes's theory led them to emphasize the role played by money, interest, and liquidity preference. Unfortunately, their critical analysis of general equilibrium was not matched by a unified and consistent proposal of an alternative theory based on a new explanation of Keynes's intuitions. Despite the importance of some of their criticisms – of which Robinson's 1953–54 argument about the heterogeneity of capital goods is a significant example – they failed to convince their fellow economists about the necessity to switch from an approach essentially based on individual choice and market equilibrium to a truly macroeconomic approach based on Keynes's insights. According to Robinson (1971: ix), Keynes himself was partly responsible for the neoclassical misrepresentation of his theory: 'Keynes himself began the reconstruction of the orthodox scheme that he had shattered'. Indeed, a very controversial passage of *The General Theory* is often quoted to support this view. In the final chapter of his most famous book, Keynes claims notably that

[i]f we suppose the volume of output to be given, *i.e.* to be determined by forces outside the classical scheme of thought, then there is no objection to be raised against the classical analysis of the manner in which private self-interest will determine what in particular is produced, in what proportions the factors of production will be combined to produce it, and how the value of the final product will be distributed between them.

(Keynes 1936/1946: 378–9)

Apparently, this statement endorses the neoclassical framework of analysis, and paves the way for the generalized use of the general equilibrium approach. At closer examination, however, Keynes's message may be seen as meaning that national output can be determined only through a macroeconomic analysis totally alien to general equilibrium. Once national output has been determined thereby, and only then, its distribution may be analysed by reverting to the microeconomic principles of neoclassical theory. On this interpretation, Keynes's statement sounds like an attempt at clarifying the relationship between macroeconomics and microeconomics, where the former is given a logical priority and

is based on its own macroeconomic foundations. Neither Robinson nor Shackle nor any of their followers interpreted this way Keynes's statement, which was considered as evidence of Keynes's inability to avoid falling prey to old ideas, 'which ramify, for those brought up as most of us have been, into every corner of our minds' (ibid.: viii). As regrettable as this may be, Keynes was blamed as being the first to endorse a neoclassical revision of his own theory, and the eventual incorporation of its original concepts within the framework of general equilibrium analysis.

The response of numerous, mainly American Keynesian economists attempting to meet the challenge of neoclassical analysis was to take over Hicks's IS–LM scheme on the assumption of wage and price rigidities. Modigliani (1944), Klein (1947), Samuelson (1948), and Hansen (1953) are the leading authors who contributed to the development of this approach, which led to the neoclassical synthesis of Keynes's analysis. Defined as the orthodox interpretation of Keynes by Snowden et al. (1994), this neoclassical reformulation of the message conveyed by Keynes's *General Theory* relies heavily on Hicks's contribution, so much so that a refusal of his IS–LM model would necessarily imply that of the Keynesian neoclassical synthesis. Let us therefore assess Hicks's model critically.

IS–LM: A critical assessment

Even though his neoclassical interpretation of Keynes's theory has been traditionally treated as the canonical reference, Hicks was neither the first nor the only one to recast *The General Theory* as a system of simultaneous equations. As Young (1987) points out, both Harrod (1937) and Meade (1937) presented, at the same time as Hicks, simplified models of *The General Theory* similar to the model proposed by the Oxian economist, and it can even be shown that two of Keynes's students, Reddaway (1936) and Champernowne (1936), had already elaborated a similar set of equations few months earlier. Yet, Hicks (1937) was the only one to combine 'the equilibrium conditions for the demand side of the economy into a single diagram' (Dimand 1988: 660). If we consider, furthermore, that Keynes himself endorsed this interpretation – when he wrote that 'I find it very interesting and really have next to nothing to say by way of criticism' (Keynes 1973a: 79) – it is not difficult to understand why Hicks's IS–LM model was preferred to the others.

Be that as it may, our main concern does not lie in investigating the origin of the IS–LM model, but rather in analysing its conceptual implications. What can be noted from the outset is that the IS–LM model is

an attempt to interpret Keynes's theory in terms of equilibrium, and on the assumption that the conditions of stock and flow equilibrium can be derived separately. The choice of the expression 'neoclassical synthesis' is not arbitrary. It clearly shows that Hicks's IS–LM restatement is the attempt to incorporate Keynes's contribution within the theoretical framework of general equilibrium analysis. It is true, of course, that Keynesian economists do not share the neoclassical belief that markets clear, that prices and wages adjust quickly to any disturbance, and that monetary shocks do not really matter (at least in the long run). Yet, it is also a well-documented fact that they agree with their neoclassical colleagues 'upon two methodological premises: that macroeconomics should be grounded in microeconomic principles, and that understanding macroeconomic behavior requires the construction of a (simple) general equilibrium model' (Greenwald and Stiglitz 1993: 24). Hicks's IS–LM interpretation of Keynes's *General Theory* is entirely consistent with these two methodological premises. Let us elaborate on them.

The IS–LM diagram is built by drawing the IS curve representing the set of equilibrium points between saving and investment, and the LM curve representing the set of equilibrium points between the supply of and the demand for money in regard of the same variables, namely income and interest rate (Figure 5.1). The IS curve is derived from the consumption function, the investment function, and the condition of equilibrium defined by the equality between saving and investment. Since both saving (consumption) and investment are considered as expenditures, the IS schedule is conceived of as a locus of flow equilibria or, alternatively, as a locus of real equilibria, because saving and investment are closely related to produced output. On the other hand,

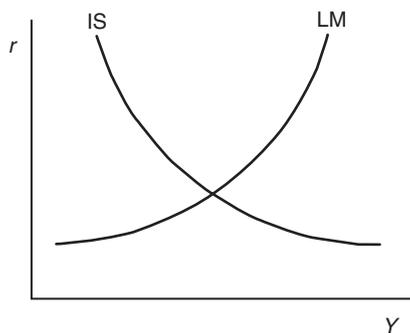


Figure 5.1 The IS–LM diagram

the LM curve is obtained by reference to a demand for money function (Keynes's liquidity preference), a supply of money function, and the equilibrium condition for the money market. Consistent with the widely shared belief that money is an asset, Hicks considers the LM schedule as a locus of stock equilibria or, when compared to IS, as a locus of nominal equilibria.

Although these stock and flow equilibria are determined separately, they can be represented in the same diagram, because they are both assumed to be functions of income and the interest rate. The IS curve is said to represent the 'combinations of interest rate and income associated with equilibrium in the goods market' (Snowdon et al. 1994: 90). Since it is assumed that investment is inversely related to the interest rate, the IS curve slopes downwards and is more or less steep according to the interest elasticity of investment and the value of the relevant multiplier. The combinations of income and interest rate that define equilibrium in the money market are represented by the LM curve. While the supply of money is assumed to be exogenously given, the demand for money is considered to be positively related to income and negatively related to the interest rate, which explains why the LM curve is sloping upwards. The income and interest elasticities of the demand for money determine the steepness of the LM curve. Finally, the intersection of the IS and LM curves determines the point where equilibrium in the goods and money markets is attained: it 'represents the *only* value of the rate of interest and income which is consistent with equilibrium in both markets' (ibid.: 94).

Two questions may be asked at this stage, namely whether the IS-LM model is faithful to Keynes's fundamental insights, and whether it is consistent with the principles of macroeconomics as determined in a monetary economy. These two questions are closely related given that Keynes is considered to be the founding father of modern macroeconomics, and that his *General Theory* is an explicit attempt to provide for a rigorous monetary theory of production. Leaving aside the problem of reading Keynes's mind, we will fuse the two questions above by assuming that the principles established by modern monetary macroeconomics are the logical outcome of Keynes's insights. Therefore, we will investigate the validity of the IS-LM model on logical grounds as an issue separable from any dispute about its compatibility with a specific interpretation of Keynes's analysis.

Let us start by examining the IS curve, which is construed on the assumption that saving and investment are at equality only for a series of combinations of income and interest rate. Equilibrium is the concept

that is applied in this framework, and the IS curve is defined as the set of equilibrium points between S and I . It suffices to recall here the analysis of Keynes's identities presented in Chapter 2 to realize that Hicks's interpretation is clearly at odds with the message conveyed by *The General Theory*, more specifically, with what is argued by its author in Chapter 7 on 'The meaning of saving and investment further considered'. As a matter of fact, modern monetary macroeconomics confirms Keynes's analysis: the identity between saving and investment is the only relationship compatible with the fact that both saving and investment are interconnected expenditures. Indeed, the identity between S and I is a direct consequence of production being itself an expenditure establishing an identity between currently produced output (global supply) and current income (global demand). In a monetary economy, where bank money obeys the rule of double-entry bookkeeping, it is logically impossible for (realized) saving and investment to differ for the simple reason that they are the two faces of the same reality: investment is the expenditure of that part of current income that is not spent on the commodity market, that is, an expenditure defining saving.

The logical consequence of Keynes's analysis is that, for whatever level of income and interest rate, saving and investment are always necessarily equal. This conclusion is certainly in contrast with the claim that S and I are equal only for the equilibrium value of income, and thus seems at variance with what was maintained by Keynes himself. In reality, an inconsistency within Keynes's own analysis would exist only if it could be possible to show beyond any possible doubt that for the author of *The General Theory* there may exist non-equilibrium levels of realized income. Once again, the problem is not an exegetical one. We are not interested in attempting to establish what Keynes's real thought was. Our concern is merely to show that his identities are not necessarily contradicted by his use of the expression 'equilibrium income level'. In fact, income is the result of production, and once production has taken place no difference can appear either between global supply and global demand or between saving and investment. This means that every produced income is an equilibrium income. To be sure, before production actually occurs, that is to say, when the latter is merely virtual, planned saving may differ from planned investment, or planned supply may be different from desired demand. However, the passage from virtual to actual magnitudes marks the passage from equilibrium to identities. Perhaps the distinction between virtual and actual magnitudes was not on Keynes's mind, or perhaps he was struggling to reconcile two incompatible conceptions of his identities. Whatever is the case, it does not alter

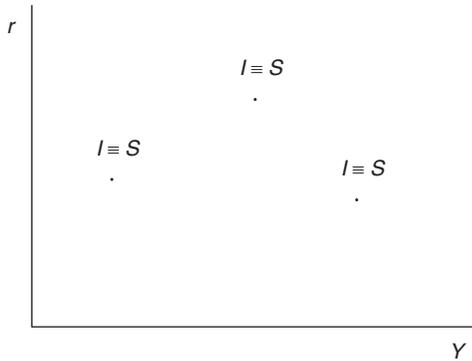


Figure 5.2 The relationship between S and I

the fact that Keynes has devoted part of his *General Theory* to the defence of the identity between S and I , and this identity is the only alternative compatible with the macroeconomic analysis of a monetary economy based on the use of bank money.

Eventually, the IS curve represents a state of affairs that has nothing to do with the reality of a monetary economy. As every level of produced income implies the identity between S and I , the IS curve has to be replaced by a point. Graphically, we could show the particular relationship existing between income, interest rate, saving, and investment as in Figure 5.2. Every point defining the identity between saving and investment is independent of the others, and represents the specific value of IS corresponding to a given production, that is, to a given amount of currently produced income. If S and I are taken to define macroeconomic saving and macroeconomic investment, every point on the graphed curve stands for the amount of macroeconomic saving and investment effectively realized for any given level of produced output.

Let us now consider the LM curve.

Originally called LL by Hicks (1937/1982), the LM curve represents the set of points for which the demand for money equals the supply of money. It is a function of income and interest rate, and is upward sloping 'since an increase in income tends to raise the demand for money, and an increase in the rate of interest tends to lower it' (Hicks 1937/1982: 153). Hicks's implicit assumption is that money is a stock whose supply is exogenously determined by monetary authorities, while its demand depends essentially on the 'transaction motive'.

In this regard, a number of critical considerations are justified with a view of further clarifying the conceptual distinction between money and income.

Both neoclassical and Keynesian economists do not draw any clear distinction between these two concepts. It is true that money is usually understood and defined as a means of payment and a store of value issued by the central bank, whereas income is due to production. However, money is also considered and treated as a positive asset entering into relative exchange with produced output. Hence, income is identified with the amount of money any given economic agent can derive from their sale of produced goods and services. The absence of an analytical distinction between money and income hinders the correct analysis of the LM curve. Does it represent the equilibrium between the supply of and demand for money, or an equilibrium between the supply of and demand for income? The question is not rhetorical, and it is worth considering separately two possible answers to it.

If the LM curve is related to money income, what has to be established is whether a disequilibrium is conceivable between the amounts of income supplied and demanded by the economy. The answer in line with Keynes's identities is clearly that no such a difference can exist in the realm of realized magnitudes. The identity between global supply and global demand leaves no room for intellectual speculation: production creates the totality of macroeconomic income, and, since income is formed as a bank deposit and banks comply with the principle of double-entry bookkeeping, the totality of income is lent and consequently demanded by some economic agents. It is because supply creates its own demand, that no adjustment occurs between the supply of and demand for money income. As a direct consequence, when related to money income, the LM curve is reduced to a single point for any given production.

Next is the question of what would the LM curve look like if supply and demand are related to money as such instead than to money income. The answer follows straight from the very nature of bank money (an instantaneous flow defined as a debit-credit of the payer and a credit-debit of the payee). If money were a net asset exogenously created and offered to the economy, its supply and demand could be distinct magnitudes whose equilibrium might be reached through interest rate fluctuations. Metaphysics apart, however, it is meaningless to look for money coming into existence with a positive or intrinsic value. The very idea that money is by nature a *stock* is mystifying if not preposterous. How can a bank issue, that is to say create, a

positive asset out of nothing? This is also at odds with double-entry bookkeeping.

The *flow* nature of money is such that it can only circulate instantaneously from and to its issuing bank. During its circular flow, money is therefore simultaneously created and destroyed by banks on each economic agent involved in a payment. Through double-entry bookkeeping, banks provide free of cost the numerical vehicle required to convey payments. What is true for money as the vehicle for payments is obviously not true for income, which cannot be produced cost-free by banks. Yet, the problem we are concerned with here relates to money as such, and to the possibility of representing with an upward-sloping curve the relationship between the supply of and demand for money (LM curve). Since money is an *asset-liability* issued by banks in an instantaneous circular movement implying both its creation and its destruction, its supply and demand are always necessarily identical. Whatever the number of money units needed to carry out a payment, banks can satisfy economic agents' demand for money by supplying the required amount of it through an emission that allows them to instantaneously recover the number of money units issued in the payment. As a consequence, the LM curve shrinks to a single point for each single payment as well as for the whole series of payments within any given period of time.

Hence, whether it is related to money income or to money as such, the LM curve is not a curve at all, and its diagram is reduced to a single point for any given level of production (Figure 5.3).

In conclusion, the IS-LM diagram is but the product of an entirely misconceived interpretation of Keynes's theory, where its logical

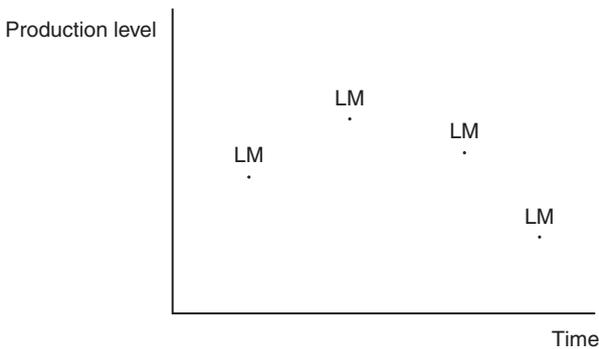


Figure 5.3 The identity between *L* and *M*

identities are illegitimately transformed into equilibrium conditions and heterogeneous variables are arbitrarily inserted into a common set of coordinates. As recognized by Hicks (1976/1982: 290) himself, the IS–LM analysis ‘reduces the *General Theory* to equilibrium economics; it is not really *in time*’. In fact, Hicks’s interpretation deprives Keynes’s theory of originality, while being incapable to provide a viable alternative. Hicks is right in identifying the shortcomings of his interpretation with the failure of general equilibrium analysis to account for the role played by time. What the Oxian economist is unaware of is that instantaneous events may define a finite period of time, so that time can be incorporated into economic analysis without considering economic magnitudes as continuous or discontinuous functions of time. Indeed, whatever the period of chronological time a given output refers to, production, consumption, saving, and investment are instantaneous events whose dimension is quantum time. As intuited by Keynes, the relationship between *S* and *I* is an identity, which can only be represented by a point in chronological time but whose quantum-time dimension is positive, because both *S* and *I* define the same amount of income and income is itself defined by production, an instantaneous event that quantizes time (Schmitt 1984a). Hicks’s diagram refers to chronological time, so that the IS curve necessarily reduces to a point, as does the LM curve.

The emission of money takes place in a point in time. Its role is to convey payments and it shares with them their instantaneity. Now, the demand for and supply of money income is unavoidably linked to production, and production is an instantaneous event in economic terms defining a well-specified payment, namely the payment of wages. Production is at the origin of money income, and it is the money income formed by production that defines both the supply of and demand for money. This clearly means that supply of and demand for money are the twin results of a single, instantaneous transaction. The necessary consequences are that *L* (the demand for money) can never be different from *M* (the supply of money), and that the LM curve is inscribed in quantum time. The same occurs for the IS curve, if it is referred to actual saving and investment. At the macroeconomic level the determination of saving coincides indeed with that of investment. Whenever a macroeconomic investment occurs, it defines the formation of a positive fixed capital, and fixed capital is nothing other than an income definitively subtracted from consumption, that is, a macroeconomic saving. On the contrary, if the IS curve is referred, *ex ante*, to virtual saving and investment, it is not reduced to a point. However, in this case IS is inscribed in continuous time, and every attempt to introduce IS and LM

curves into the same diagram is doomed to failure: 'even if the IS curve were to be inscribed in the continuum, it would never meet the LM curve, which pertains to an entirely different space' (Schmitt 1984a: 569, our translation).

Entirely at odds with an analysis respectful of Keynes's identities, the IS-LM model is responsible for the widespread acceptance of neoclassical Keynesianism in the 1960s as well as for the development of New Keynesian economics. Let us briefly turn our attention to this school of thought.

New Keynesian economics

The unexpected success of new classical economics forced Keynesianism to choose between two alternatives: (1) go back to Keynes's original intuitions and work out a truly macroeconomic approach to macroeconomics, or (2) take over the neoclassical challenge and ground Keynesianism deeper into microeconomic analysis. A number of well-known economists, such as Akerlof, Blanchard, Greenwald, Mankiw, Phelps, Romer, Stiglitz, Summers, and many others, chose the second alternative and thereby developed what has since been called New Keynesian economics.

New Keynesians aim to provide adequate microeconomic foundations to the thesis that the labour market is far from being in a continuous state of Walrasian equilibrium and that monetary disturbances have real effects. They emphasize the role played by changes on the demand side and analyse the way these changes affect the level of employment in an attempt to show that unemployment is an important phenomenon that can be explained by aggregate disturbances arising from both exogenous and endogenous shocks, which are often greatly amplified by the presence of frictions and market imperfections giving rise to substantial fluctuations in the level of real variables.

Since Keynesian economics had failed to substantiate the claim that price adjustments are imperfect and that nominal frictions may have a disruptive impact on the whole economy, New Keynesians take over the challenge to provide sound microeconomic foundations to these claims, notably to the idea that 'small frictions can cause nominal disturbances to have large effects on aggregate economic activity' (Romer 1993: 8). They do so by arguing that, owing to the existence of small barriers to nominal adjustments, there are price rigidities that cause movements in aggregate demand to have large effects on economic stability. Nominal frictions are believed to have a larger impact if they are

introduced into a model assuming real rigidities to be low (Romer 1993), and the factors considered as potential sources of real rigidities go from asymmetric information to the elasticity of demand on the goods market, from the elasticity of labour supply to that of real wages. Imperfect price adjustments are said to cause macroeconomic changes in the presence of nominal disturbances and aggregate demand shocks, and New Keynesians' main goal is to provide substantial microeconomic evidence confirming the role played by market imperfections and nominal fluctuations in a world where perfect competition and general equilibrium are figments of the neoclassical imagination. According to Gordon (1990: 1117), stickiness in price and wage adjustments emerges from microeconomic elements such as imperfect competition, asymmetric information and capital market rigidities, which 'remove any incentive for individual agents to focus on nominal demand in making their own price-setting decisions'.

In fact, New Keynesian economists are not unanimous in considering nominal price and wage rigidities as the essential cause of macroeconomic disturbances. As stressed by Greenwald and Stiglitz (1993: 25), another 'strand of New Keynesian literature explores another path suggested by Keynes: that increased flexibility of wages and prices might exacerbate the economy's downturn'.

What both approaches have in common is the belief that monetary and non-monetary shocks are the rule and have an impact on macroeconomic variables that is magnified by market imperfections and asymmetric information. Whether they emphasize nominal or real price rigidities, risk costs or costs of price adjustments, the need for a structural reform reducing nominal price and wage rigidity or reducing financing imbalances, economists advocating a New Keynesian analysis agree on the need to start by investigating firms' and consumers' *behaviour* in order to lay their models on microeconomic foundations. None of them accepts the neoclassical market-clearing axiom nor the existence of representative agents, yet they all seem to agree 'on two methodological premises. That macroeconomics should be grounded in microeconomic principles, and that understanding macroeconomic behaviour requires the construction of a (simple) general equilibrium model' (*ibid.*: 24).

As a matter of fact, New Keynesians have taken over most of the assumptions of neoclassical and new classical analysis. Besides the primacy given to microeconomic foundations, they usually accept Lucas's rational expectations hypothesis as well as the general equilibrium theoretical framework: 'macroeconomics today is solidly grounded

in a general equilibrium structure' (Blanchard 2000: 36). The new neoclassical synthesis as described by Goodfriend and King (1997) in a working paper of the Federal Reserve Bank of Richmond incorporates Keynesian and New Keynesian elements on one side, and neoclassical, new classical, monetarist, and real business cycle elements on the other.

Methodologically, the new synthesis involves the systematic application of intertemporal optimization and rational expectations as stressed by Robert Lucas. In the synthesis, these ideas are applied to the pricing and output decisions at the heart of Keynesian models, new and old, as well as to the consumption, investment, and factor supply decisions that are at the heart of classical and RBC [real business cycle] models.

(Goodfriend and King 1997: 232)

Mainstream economists are therefore increasingly involved in the construction of macroeconomic models combining intertemporal optimization and rational expectations with imperfect competition and price stickiness. They believe that aggregate demand plays a central role in the determination of economic activity in the short run, and that economic fluctuations are influenced by monetary policy, as well as by productivity and fiscal shocks. As stated by Woodford (2008), a wide agreement exists today among macroeconomists as to the need to:

1. develop 'models with coherent *intertemporal general-equilibrium foundations*' (p. 3) incorporating 'a variety of types of adjustment frictions that allow these models to provide fairly realistic representations of both shorter-run and longer-run responses to economic disturbances' (p. 4);
2. 'base quantitative policy analysis on *econometrically validated structural models*' (p. 5);
3. '*model expectations as endogenous*' (p. 8);
4. accept that '*real disturbances* are an important source of economic fluctuations' (p. 9); and
5. recognize that monetary policy matters.

Divergences remain, yet they are limited to the choice between alternative dynamic stochastic general equilibrium (DSGE) models. In particular, strong disagreement exists between those authors who maintain that models should be kept as simple as possible (Chari and Kehoe 2007; Galí et al. 2007), and those who, in an attempt to fit optimally

macroeconomic aggregates, develop more and more complex models (as in Smets and Wouters 2007). Finally, let us note that, according to Blanchard (2009: 224), DSGE models are still far from being satisfactory, because '[r]econciling the theory with the data has led to a lot of unconvincing reverse engineering'. In his paper on the state of macroeconomics, Blanchard (2009) maintains that the introduction of assumptions such as that of external habit formation, convex costs of changing investment, and backward indexation of prices undermines the validity of DSGE models. 'Because their introduction can then be blamed on others, these assumptions have often become standard, passed on from model to model with little discussion. This way of proceeding is clearly wrongheaded' (Blanchard 2009: 224). This leads Blanchard to plead for the 'rehabilitation of partial equilibrium modelling in macroeconomics' (ibid.: 225) as well as for the 'relegalization of shortcuts and simple models' (ibid.: 226).

All in all, agreement among mainstream macroeconomists is considerably more important than disagreement, and the success of the new neoclassical synthesis clearly shows how strong the influence exerted by neoclassical analysis has been, and how far the New Keynesian approach is from Keynes's search for a macroeconomic theory with deep and solid macroeconomic roots. No one will ever know whether Keynes would have endorsed some of the New Keynesian claims or not. It is indisputable, in fact, that Keynes himself provides elements for justifying the Keynesian and New Keynesian drift away from his key insights while advocating, at the same time, the need for a new macroeconomics founded on the identity between Y and $C+I$. What really matters is not whether New Keynesian economics is faithful to Keynes's own thought or not, but whether it is consistent with facts, that is, with the very nature of economic reality.

Our critical appraisal of New Keynesianism will therefore not be focused on its alternative models, nor will it refer to the trustworthiness of their assumptions. The need to found macroeconomics on microeconomics or not is not a question of choosing between two equally valid alternatives. Likewise, the use of general or partial equilibrium models is not a matter of methodological preference or technical convenience. The apposite question to be asked is whether macroeconomics obeys its own logical laws or not, if it results from a mere aggregation of microeconomic variables or is a distinct field of inquiry called for by the reality of our economies. Therefore the answer must rest on doing justice to the very nature of our economic system, which is a *monetary* economy of production and exchange. This brings us to the core

of the problem, that is, to the distinction between *relative* and *absolute* exchanges. Indeed, production itself can be analysed as an exchange, and so can consumption, lending, and any other conceivable economic transaction. It is therefore of paramount importance to establish what absolute exchange means, and how it differs and connects to relative exchange.

The determinant factor for understanding the two kinds of exchange is money. If money were an asset, as claimed by neoclassical as well as New Keynesian economists, its exchange with productive services could be thought to be of a relative kind. It is true that, even in this case, the existence of relative exchanges would still have to be formally proved, since it is far from self-evident that production prices can be determined through the relative exchange of money and productive services. Alas, the logical proof of the indeterminacy of relative prices (see Schmitt 2012) is highly conceptual and counter-intuitive, and would discourage most readers not well acquainted with the problem. However, the analysis of bank money comes to the rescue for these readers. Indeed, the claim that money can be issued by banks as a positive asset is completely misconceived. As shown in Chapter 1, money is a mere numerical form and the exchange defining production does not take place between two distinct assets (money and output). The payment of wages is the only transaction allowing money to give output a numerical form and to transform output into the real content of money. Production, that is, the exchange through which money acquires a real content and output takes the form of income, is an absolute exchange: physical output is literally 'changed' into a sum of income, which represents its economic expression. In other words, production is an absolute exchange that gives rise to a double-faced object, namely income, which defines at the same time the purchasing power required to allow for the final purchase of produced output and the value of this same output.

The outcome of production is not a physical output on one side and a sum of income on the other side. As clearly stated by Smith (1776/1991) and confirmed by national accounting, the outcome of a production generating an income of, say, x million US dollars, is worth x millions and not twice this amount. The value of produced output does not add up to that of income, and this is so because currently produced output is the object of current income's purchasing power. This means that macroeconomic supply and demand are the twin results of a unique transaction, or the twin aspects of one and the same reality.

The logical identity between macroeconomic supply and demand is not a matter of aggregation of microeconomic relationships, and

has nothing to do with individual behaviour. Irrespective of the number of individuals involved and of their decisions, production is a macroeconomic event because it modifies the situation of the economic system considered as a whole, and it does so by giving birth to an economic output, that is, to a physical output cast in a numerical mould. The twin results of production are therefore a positive amount of income (defining a macroeconomic demand) and its real content (the produced output defining macroeconomic supply). Keynes's identity between Y and $C + I$ is a fundamental macroeconomic law deriving from monetary production, and not a condition of equilibrium as assumed by Keynesians and New Keynesians. By taking over Hicks's interpretation of *The General Theory* as well as the neoclassical concept of general equilibrium (albeit considered as imperfect or non-Walrasian), New Keynesians turn their back to Keynes's identities and to the possibility to found macroeconomics on solid macroeconomic ground. Their analysis differs only marginally from that advocated by other schools stemming from the general equilibrium approach and, like them, is at odds with the reality of our monetary economies of production. It is a fact, and not a theoretical assumption, that money is bank money and that it is issued according to the principle of double-entry bookkeeping. It is also a fact that banks cannot create wealth and that money acquires a positive value only through production. Macroeconomic laws have their origin in these facts as well as in the discovery that it is only through the payment of wages that physical output is transformed into an economic product. New Keynesians work out their models according to a set of assumptions that do not respect these facts and fail to recognize the monetary nature of economic production. Their analysis is confined to the description of an imaginary world where money is an asset created by banks, economic agents must behave rationally, monetary and real variables are dichotomous magnitudes, and prices are relative: a world that has little to do with the one in which we live.

Does post-Keynesian analysis fare better? Is it in particular capable to account for the macroeconomic foundations of macroeconomics? Let us address these questions in the remainder of this chapter.

Post-Keynesian economics

Post-Keynesian economics aims at representing a radically different alternative to the more orthodox schools of economic thought including Old and New Keynesians, who are 'bastard Keynesians' according to Robinson (1962: 690). Despite encompassing a number of different and

sometimes contradicting lines of thought, the post-Keynesian approach to economic analysis unifies all of them in their attempt to reconsider and to modernize Keynes's original analysis, elaborating upon it in order to explain a variety of contemporary issues, such as unemployment, financial instability, and crises in an open-economy framework. This unifying characteristic is however problematic, since the various strands of post-Keynesian thought consider and understand differently the major publications of Keynes, that is, his *Treatise on Money* and *The General Theory*. The issues at stake concern in particular the nature and role of money in our capitalist economies. These issues are crucial, indeed, for the working of a monetary production economy. Depending on their treatment by post-Keynesian scholars, they bring the latter closer to or farther from more orthodox economic thought.

Understanding money's nature allows one to grasp the essential function of banks in payment systems from a macroeconomic perspective. Banks are actually special, for they provide the numerical instrument to measure produced output and record it in the form of bank deposits as explained in Chapter 1. The investigation of post-Keynesian economics must therefore start from money and banking, before addressing a variety of ancillary issues such as inflation, unemployment, financial instability and crises. In fact, without banks, money would not exist as a means of final payment, making thus impossible for any economic system to operate, because the latter necessitates a set of relations based on payment finality, that is, the process of debts settlement, as a result of which the payer has no further debt to the payee, who has thus no further claims on the payer (Goodhart 1989: 26).

In fact, post-Keynesian economics is split into two camps, namely between those who consider money to be a means of payment issued by banks when they provide a credit to any kind of agents, and those who believe that money is a liquid store of value that owes its existence to fundamental uncertainty and agents' liquidity preference. Let us begin our critical investigation with the latter, because this group of post-Keynesians is closer to orthodox economic thinking than the former camp, which includes so-called monetary circuit theorists (or 'circuitists').

According to Davidson (1978), Chick (1983), and many other post-Keynesians who have been inspired by them, money owes its existence to the fact that the future is unknown and unknowable (Keynes 1936/1946). This then induces agents' liquidity preference in the form of money holding (or even hoarding, an idea that we already criticized in Chapter 2). 'In their view, money is the fundamental macroeconomic

institution for coping with the uncertainty of individual decision-making. Liquidity preference (money holding) explains why in modern economies, expenditure may fall short of income, with the result that the sales revenues of producers will not cover all production costs' (Fontana 2009: 61). Its different line of approach notwithstanding, this closely echoes the orthodox attempt at providing microeconomic foundations to macroeconomic analysis. This is so much so that, according to this post-Keynesian camp, money is an asset analogous to many others, the only distinguishing feature of it being its (higher) liquidity, which is the result of money's acceptability within the payments system: 'the feature which distinguishes money from credit is the general acceptability of deposits, as against the personal quality of credit' (Chick 2000: 131). Blurring the distinction between money (a purely numerical flow) and bank deposits (liquid stocks of purchasing power), this quotation restates the famous claim made by Minsky (1986: 228), according to whom 'everyone can create money; the problem is to get it accepted'.

Generally speaking, post-Keynesian economists consider that money's acceptance is the result of a social convention (similar to Tobin's (1963) famous analogy between money and language) and that money is itself a social relation (Ingham 1996). They substantiate this belief with the image of a debt pyramid: 'the "hierarchy of money" can be thought of as a multi-tiered pyramid where the tiers represent promises with differing degrees of acceptability' (Bell 2001: 158). Still more precisely, this 'can be envisioned as a four-tiered debt pyramid, with the debts of households, firms, banks and the state each representing a single tier' (ibid.: 159). Moving from this 'simplified hierarchy' (ibid.: 159), an increasing number of post-Keynesians (united under the self-established label of 'modern money theory') led by Wray (1998), argue that the purchasing power of money depends on the power of the State to impose a number of tax liabilities to the population, as observed in Chapter 1. Therefore, this theory 'links the state's ability to issue a currency denominated in the unit of account it has chosen [...] to a fundamental power that is directly associated with sovereign nations' (Wray 2003a: 89). This goes back to the chartalist argument that money is not a creature of the market (as argued by metallists) but a creature of the State (see Goodhart 1998), because the latter has been entitled by law 'to declare what thing should answer as money to the current money of account' (Keynes 1930/1971: 4). In light of this view, chartalists argue that government fiat money is a liability of the State, which the latter agrees to accept at State pay offices and that therefore is largely used in the payments system of the relevant country (Wray 2003a: 89–90).

This confusion between ‘State money’ and central bank money – which stems from the view that the Treasury and the central bank are like ‘husband and wife within the household’ (ibid.: 92) – induces the wrong belief that ‘the [US] government can buy anything that is for sale for dollars merely by issuing dollars’ (Wray 1998: ix). Hence, the State should act as an Employer of Last Resort (ELR), and hire therefore all those unemployed who are ready, willing, and able to work at a minimum wage determined by it (Wray 2003b). As cogently argued by Sawyer (2003: 884), however, ‘the ELR “solution” appears to be to provide employment which does not require any capacity or skill’. This ‘solution’, in fact, does not seem to consider the problem of structural unemployment, that is, the mismatch between the skills looked for by firms and the skills offered by unemployed on the labour market (see Kadmos and O’Hara 2000). Further, and more important, this ‘solution’ critically depends on the public sector to issue as much money as is necessary to have full employment (Kadmos and O’Hara 2000: 10; Sawyer 2003: 885–7). This not only contrasts with the practice of balancing the public sector’s budget that a number of countries adopted in the aftermath of the neoliberal revolution started in the 1970s, but also with the nature of money that we pointed out in the first chapter of this book. Indeed, when the State pays a minimum wage to any individual participating in the ELR programme as in the chartalists’ view, it just gives to its creditors a promise to pay in the form of financial claims (bank notes or deposits with the central bank). This payment is not final: it is a mere promise of payment finality, because it is only when the public disposes of this ‘State money’ at public pay offices, for the payment of tax liabilities or any other debt obligations, that the transactions between the State and the private sector economy are settled. To be sure, in this framework payment finality occurs by a sort of barter trade, where money acts as a medium of exchange in the best neoclassical tradition: goods, services (including labour services), and assets are bartered against fiscal obligations, with ‘State money’ intervening as an ‘intermediary asset’. As a result, the economic system advocated by proponents of ‘modern money theory’ is essentially a money-using barter-trade system and not really a monetary economy of production (Rossi 2007a: 16–22).

Let us turn therefore to the other post-Keynesian camp, which considers money as the means of final payment that banks issue when they provide credit to any kind of agent within a monetary economy of production, but particularly to those firms that need an ‘initial finance’ (Graziani 1990: 14) to start a new production period. In this view, the

credit provided by banks has a purchasing power *per se*, because it stems from banks, which are considered as endowed with the capacity to 'anticipate liquidity' to firms 'in order to cover the current cost of output' (ibid.: 14). According to this perspective, '[m]oney is in the nature of credit money and in modern times is represented by bank credit' (Graziani 2003: 25). Credit money is issued when a payment is made: 'At that moment, in one and the same act, money is created, the borrower becomes a debtor to the bank and the agent receiving a payment becomes the creditor of the same bank' (Graziani 1990: 11).

In fact, money and credit must be distinguished essentially, even though the emission of money is tied to credit in the sense that 'loans create deposits' within the banking system, as noted by Schumpeter (1954/1994: 1114). When a bank issues money, as a matter of fact, it has a claim on the payer that is balanced by an equivalent claim that the payee has on the bank. Now, the claim owned by the payee in the form of a bank deposit defines his credit against the issuing bank. This, nevertheless, does not mean that the bank lends the number of money units that it issues in a payment. In fact, the underlying loan concerns the two agents involved in the payment: the payee grants a credit to the payer through the bank (or banking system) acting as intermediary, even though both agents involved might not be aware of this financial intermediation (see Gnos 1998). As explained in Chapter 1, the bank is neither a net creditor nor a net debtor of the whole economy when it issues any number of money units, since it is simultaneously debited and credited with the same amount.

This is enough to understand that money and credit are essentially separate concepts, and that therefore the 'monetary circuit' described by so-called 'circuitists' is, in fact, a credit circuit, which takes time to complete. Indeed, when 'circuitists' argue that the monetary circuit has a positive duration in time, because '[c]redit money is created whenever an agent spends money granted to him by a bank and is destroyed whenever a bank credit is repaid' (Graziani 2003: 25), they are in fact referring to the circuit of bank credit, that is to say, to the period of chronological time separating the formation of income on the factor market from its final expenditure on the market for produced goods and services (see Chapter 2). As done by those post-Keynesians who claim that money's existence is essentially the result of agents' liquidity preference in a fundamentally uncertain world, 'circuitists' adopt an analogous (microeconomically founded) line of argument, when they imagine that '[l]iquidity preference represents a problem for the closure of the system [that is to say, the monetary circuit]' (Rochon 1999: 35).

In reality, it is wrong to maintain that an increase in bank deposits 'poses a problem for the closure of the monetary circuit [and that] hoarded saving represents a leakage' (ibid.: 35). As a matter of fact, owing to its bookkeeping nature, no bank deposit can leak out of the banking system where it is recorded until it is spent on the market for produced goods and services (and thereby destroyed). What Rochon labels 'hoarded savings' are in reality immediately lent – by the bank where they are recorded on the liabilities side of its balance sheet – to those businesses that can finance their factor costs neither on the product market nor on the financial market (through the sale of securities). In light of the analysis presented in Chapter 2, we can easily notice that the explanation of 'monetary circulation' put forward by 'circuitists' is not up to the task of explaining the nature of money, as it is based on the simplistic view that the 'flow of money' is a stock moving around within the economic system (as stated by Robertson (1937: 29) with his famous expression 'money on the wing'). Again, one is thus trapped by a physical (albeit not material) conception of money, as in the more orthodox analyses discussed above, where agents' forms of behaviour are (implicitly or explicitly) crucial to explain the nature and role of money in a capitalist economy.

Indeed, post-Keynesian economics focuses on behavioural factors to explain not only money's existence and function(s) but also macroeconomic pathologies like inflation, (involuntary) unemployment, financial and economic crises. Inflation, which the post-Keynesian community considers as defined by a prolonged increase in the price level – thus conforming to the orthodox superficial understanding of this phenomenon – is therefore explained by a conflict over income distribution between firms and workers, whose conflicting claims on income increase prices on the market for produced goods and services (Rowthorn 1977). As Davidson (1991: 92) puts it, '[i]nflation is always and everywhere a symptom of the struggle over the distribution of income'. With regard to involuntary unemployment, post-Keynesians have generally explained it by a lack of 'aggregate demand' (another micro-founded notion based on orthodox thought): linking it to agents' liquidity preference and to 'animal spirits' of entrepreneurs, post-Keynesians maintain that the behaviour of the public sector could (and should) make sure to close the gap between aggregate demand and supply at the full-employment level. Agents' propensities to consume or to invest are thus to be increased by a surge in public spending, particularly when the economy is in a recession or even in a depression (see Davidson 1998; Stockhammer 2011). It should not come as a surprise,

then, that post-Keynesian authors explain economic and financial crises also with regard to behavioural factors, often referring to Minsky and particularly his now famous 'financial instability hypothesis' (Minsky 1982: 95): a period of steady economic growth induces riskier forms of behaviour by financial-market operators. The debt structure of the economic system becomes problematical, as agents reduce the security margin on debts in light of the apparent stability of that system. When the prevailing financial conditions worsen, however, the whole system becomes unstable, leading to a crisis (see Kregel 2008; Wray 2009; Dymski 2011). In this permanent tension between agents' behaviour and their implications at 'systemic' level lies both the force and the weakness of post-Keynesian economics: its force can be observed and developed when post-Keynesian authors consider that the economic system as a 'whole is different from the sum of the parts' (Ormerod 1994: 91). This is actually the novelty, then and now, of Keynes's original message (different from that of self-declared 'Keynesians' of any sorts), as argued throughout this book. However, when post-Keynesians reduce all 'systemic' outcomes to agents' behaviour, including different institutions (such as central banks, financial supervisory authorities, and the general government sector), their approach to macroeconomic issues suffers from the same fundamental weakness that vitiates more orthodox lines of thought, that is, the willingness to explain the (orderly as well as the disorderly) working of an economic system with regard to its alleged microeconomic foundations. King (2012: 5) points this out cogently when he argues that 'many Post Keynesians have been strong critics of the [microfoundations] dogma, but a surprisingly large proportion [of them] have been unclear, inconsistent or just plain confused on the issue.' This is particularly the case with regard to the post-Keynesian explanations of the global financial crisis that burst in 2008 after the failure of the investment bank Lehman Brothers in the United States (Rossi 2010). Let us turn therefore to a truly macroeconomic analysis of crises, which will show both the shortcomings of microeconomically founded explanations of them and the explanatory power of macro-founded macroeconomics in this regard.

6

Economic Crises and Their Relationship to Global Supply and Global Demand

Does it make sense to analyse crises in terms of global supply and demand? Can they be explained consistently with the necessary equality between supply and demand required by modern macroeconomic analysis? These are the crucial questions we will address in this chapter. Starting from Say's law and Keynes's logical identity between Y and $C+I$, we will first investigate the problem of whether or not the insurgence of an economic crisis entails their rejection. Indeed, the possibility of reconciling a situation of disequilibrium with the identity of global supply and demand seems very remote if not altogether inexistent. On the other hand, however, quantum macroeconomics provides clear logical evidence that the identity between global supply and demand is at the heart of economics. This can only mean that, eventually, economic crises will have to be explained without denying this identity. This is not what the followers of mainstream economics claim. Both neoclassical economists (whether advocates of the New Classical or of the real business cycle approach) and Keynesian economists (whether members of the New Keynesian or the post-Keynesian school) believe in some kind of general equilibrium framework and ascribe the outbreak of economic crises to factors affecting either the supply side or the demand side of their models. The opposition between the theoretical frameworks of quantum macroeconomics and mainstream economics as to the role of identities and conditions of equilibrium reveals the existence of an unbridgeable gulf between these approaches. Yet, it is a matter of factual evidence that

- 1) production is at the origin of both a supply (output) and a demand (income) defining the two identical aspects of one and the same economic magnitude; and

2) a crisis is a pathological state of the economy in which supply and demand are no longer at equality.

Our initial problem is therefore still with us. We still need to wonder whether economic crises can be explained by simultaneously

- respecting the identity between global supply and global demand, and
- allowing for a numerical difference between them.

This chapter aims at clarifying the terms of this problem and paves the ground for the answer finally revealed in Chapter 7.

Economic crises and the identity between global supply and global demand

The belief that global demand falling short of global supply is an important cause of economic crises has its origins in the analysis of mercantilists. ‘The first suggestion that there might be deeper causes to these breakdowns, causes which are inherent in the economic process, are indeed to be found in the “mercantilist” literature, mainly in connection with the ideas that were later on worked up into the various underconsumption theories’ (Schumpeter 1954/1994: 738). Yet, it may reasonably be argued that the first rigorous attempt to show that global demand may not be identical to global supply is to be found in Marx’s analysis of surplus-value. Indeed, Marx’s investigation of the way surplus-value is formed is closely related to the idea that consumption may leave some of produced output unsold and lead thereby to overproduction. Let us recall that, according to Marx (1867/1976), exchange-value is determined by labour alone and that, even though exchanges can only take place between equivalents, a surplus-value is nevertheless formed (and explained) by distinguishing between labour and labour-power. Firms pay workers their due and yet manage to obtain a surplus-value since wages are equivalent to what is sold by workers: their labour-power, and not the total labour time firms obtain from its use. Surplus-value is thus linked to the possibility for firms to exploit their workers so that a positive difference is formed between the value of total labour time and that of workers’ labour-power. The problem that arises at this stage of Marx’s analysis is known as the ‘realization problem’. Its statement is simple: the output obtained by firms at zero costs has to be sold. Now, the only income available is that formed through the payment of wages. This means that only that part of produced output whose value is equal to

that of workers' labour-power can be sold on the product market: goods corresponding to firms' surplus-value are bound to remain unsold.

At this stage the question that must be asked is whether the realization problem concerns the nature of the capitalist system or Marx's own analysis. The answer is straightforward: it is Marx's theory of profits, incapable of explaining their monetary formation, which is inadequate. If positive profits are to be explained by the theory, it is not enough to show how a positive surplus-value may be formed, but it must also be shown how surplus-value may be transformed into an equivalent amount of money income. The crucial shortcoming of Marx's analysis is that the only income it can account for is that formed through the payment of wages. Yet, global demand cannot be shown to be lesser than global supply by preventing demand to be formed in the first place. The pathological working of capitalism cannot be attributed to Marx's failure to provide a satisfactory analysis of profit, especially to his failure to explain the existence of profit in its monetary form.

It is not the existence of positive profits that is at stake. Marx is well aware of the fact that capital accumulation is a process requiring the presence of monetary profit. His aim is to show that capitalism is a system in which capital accumulation is bound to give rise to increasingly disruptive crises that will finally lead to its self-destruction. However, these crises cannot arise directly from production. Marx's error consists in assuming that profits are formed irrespective of the possibility to sell the totality of output. In reality, no meaningful distinction can be drawn between labour and labour-power, and it is erroneous to define the payment of wages as a *relative* exchange between two equivalent terms, labour-power on one side and monetary wages on the other. What Marx misses is that production is an *absolute* exchange, as defined in the first chapter, and that through the payment of wages output is 'changed' into a sum of money income. It is the totality of output that is the object of this absolute exchange, and not only part of it, and this is so because wages are the remuneration of labour and not of workers' labour-power.

The rejection of Marx's distinction between labour and labour-power is the unavoidable result of the critical appraisal of his analysis of surplus-value. One of its consequences is that a shortage in global demand cannot derive from profit formation, which leaves Marx's theory of crises at an impasse. If production is an absolute exchange taking place between money wages and current output, global demand and global supply are necessarily its twin results. How is it possible, then, to attribute the cause of economic crises to a shortage in global demand?

An answer to this question seems to be provided by Keynes's analysis of unemployment and by his concept of effective demand.

In his *General Theory*, Keynes claims that firms determine the level of output and employment on the basis of the principle of effective demand.

The amount of labour N which the entrepreneurs decide to employ depends on the sum (D) of *two* quantities, namely D_1 , the amount which the community is expected to spend on consumption, and D_2 , the amount which it is expected to devote to new investment. D is what we have called above the *effective demand*.

(Keynes 1936/1946: 29)

The first idea conveyed by Keynes's concept of effective demand is simple and straightforward: firms decide the amount of output to be produced and, therefore, the level of employment, by referring to their sale expectations on the market for consumption goods and to their expected investments. The second meaning that can be attributed to Keynes's effective demand is less obvious and relates to the output of production. Once firms have taken their final decisions and once production has effectively taken place, a demand arises that is defined by the income generated by production, and which is exerted in the purchase of consumption and investment goods, that is, Keynes's $C + I$. Whether we consider the concept of effective demand as referring to firms' expectations, *ex ante*, or to consumers' and investors' expenditures, *ex post*, it appears that the level of employment is determined in terms of output and of the income that defines it.

The question that is called for at this stage of the analysis is whether the principle of effective demand can explain the existence of a discrepancy between global supply and global demand. In order to answer this question, it is first necessary to clarify the relationship between effective demand and unemployment. When Keynes claims that effective demand determines the level of output and employment, he is opposing the neoclassical assumption that general equilibrium rules our economic systems and, through market clearing, is capable to guarantee full employment provided wages are free to fluctuate. The author of *The General Theory* disagrees, and argues instead that equilibrium is compatible with unemployment or even with overemployment. To put it shortly, he claims that the level of output determined by effective demand may well be different from the level required to reach full employment, even though it defines a perfect equilibrium, an

identity, between Y and $C + I$, that is, between global supply and global demand.

Keynes's dissociation of equilibrium from full employment calls for a clarification of the notion of unemployment. Indeed, the term 'unemployment' may stand for two partially different concepts. On one side, by unemployment one can define a situation in which the number of people able to work is greater than the number of people actually working, while the economic system is in a state of perfect equilibrium, with no inflation and no deflation. On the other side, one can define unemployment as the consequence of a pathological state of the economy, deflation, where global demand falls short of global supply. Since our main concern is the relationship between global supply, global demand, and economic crises, it is clearly the second conception of unemployment that is appropriate here. Economic crises are induced by pathological conditions altering the orderly working of our economic systems. Unemployment is one of the manifestation of economic crises, but only insofar as it is of a pathological origin and not simply one possible state characterizing an orderly system, that is to say, a system consistent with the identity between Y (global supply) and $C + I$ (global demand).

Effective demand is never scarce or in excess *per se*, but it may be insufficient to guarantee full employment. Suppose the full-employment level of output to be equal to 100 money units. If expectations were to induce firms to actually set production at the level of 90 money units, 10 per cent of workers would fail to be employed. Yet, production would define an *ex post* effective demand of 90, that is, a demand equivalent to currently produced output. The equivalence between global supply and global demand is what allows us to maintain that the economy under scrutiny would suffer from no crisis and that its unemployment would not be of a pathological nature. It is not hazardous to forecast that – under these circumstances – increasing and diversifying production would gradually absorb unemployment. If this is not what happens in today's real world, it is simply because our economies are hit by a pathological decrease in global demand, whose origin has so far been missed by mainstream economists. It is the unemployment deriving from a shortage in global demand that is worrisome, and not that consistent with Keynes's identity between Y and $C + I$. For any possible level of effective demand there is a level of employment that Keynes would have defined as consistent with equilibrium. Thus understood, the principle of effective demand has therefore nothing to do with the existence of a pathological discrepancy between global supply and global demand.

Unfortunately, the analytical distinction between *ex ante* effective demand and global demand has gone unnoticed by mainstream economists, who consider the equality between aggregate demand and supply as a condition of equilibrium. By doing so, they entirely miss the specific macroeconomic character of Keynes's analysis. In particular, they remain unaware of the existence of a relationship between macroeconomic supply and macroeconomic demand that is independent of economic agents' behaviour. Before production actually takes place, an adjustment may occur between *virtual* supply and *virtual* demand. Firms adjust their plans to expectations and take their decisions accordingly. However, once *ex ante* or virtual effective demand has played its role, production takes the lead and brings about an *ex post* or realized effective demand defined by Keynes as the sum of $C+I$. Considered *ex post*, effective demand is a macroeconomic magnitude identically equivalent to global supply, Y .

Keynes's followers pay little attention to the equivalence between Y and $C+I$, and transform what was meant to be an identity into a condition of equilibrium between aggregate supply and demand. Instead of looking for a pathology operating at the macroeconomic level, they analyse supply and demand from a microeconomic point of view and endeavour to show that demand fluctuates according to variations in economic agents' decisions. This is what happens, for example, with New Keynesian models. Indeed, even though New Keynesians are prepared to consider technological shocks as a possible relevant cause of unemployment, they make it clear that this would be the case only to the extent that 'the effects of technology shocks on employment are conditioned by the response of aggregate demand' (Galí 2012: 12). Whether the impact on aggregate demand is said to depend on the monetary policy rule in place or not, old and new advocates of Keynes's *General Theory* agree on the primacy of *aggregate* demand and believe that its variations may be due to monetary or real shocks in a framework where wage and price rigidities, imperfect competition and asymmetric information are distinctive features. In this framework, demand at the macroeconomic level is obtained by aggregating individual demands, and models 'are derived from micro foundations – that is, utility maximization by consumers-workers; value maximization by firms; rational expectations [...] – and typically estimated by Bayesian methods' (Blanchard 2009: 223).

If the analysis advocated by Keynes's followers were correct, unemployment would have two different, albeit compounded, causes. It could either be due to an expected or *ex ante* effective demand insufficient to

take production to the level of full employment or due to an (*ex post*) aggregate demand that falls short of aggregate supply. In neither of these two cases would unemployment be generated by a macroeconomic pathology affecting the economic system as a whole. Now, this would not represent a problem if it could be proven that deflation is a disequilibrium or imbalance imputable to economic agents' behaviour and that no true macroeconomic law exists. This is indeed the mainstream point of view, epitomized by the claim that aggregate demand may fail to match aggregate supply because of a change in consumers' and/or investors' propensity to consume and invest. It is thus maintained, for example, that if consumers were to systematically increase their propensity to save and firms did not increase their investment, aggregate demand would fall below aggregate supply and deflation would settle in causing a rise in unemployment. The question we have raised is therefore very clear: can a change in consumers' behaviour cause global demand to differ from global supply? In particular, can a reduction in consumers' propensity to spend induce a pathological decrease in global demand? The answer depends on the correct understanding of the term 'global demand'. Does global demand refer to the decisions taken by consumers and investors as to the amount they are going to spend for the purchase of consumption and investment goods, or is it defined by the total amount of money income available within a given economy? A positive answer to the former questions requires a positive answer to the latter. It is therefore necessary to start from a clear definition of global demand.

Let us avoid any misunderstanding from the outset. The question we have to address in this regard is not merely nominal or terminological and cannot be answered simply by arbitrarily introducing the definition best suited to fit our personal theory. The definition of global demand must result from a scientific investigation of the way income is formed and spent, and has to be consistent with the concepts of economic production and consumption. As shown in Chapter 2, production is at the origin of a positive income deriving from the association or the coupling of money and produced output. Hence, it suffices to observe that a positive demand can only be exerted by a positive income to realize that global economic demand is univocally defined and expressed by the amount of income formed by production. The idea that saving could reduce global demand if it were greater than investment is contradicted by facts, as the totality of saved income is necessarily deposited with the banking system. Double-entry bookkeeping is such that not even the smallest fraction of unspent income can abandon the banking system

where it is initially formed. As intuited by Keynes, saving and investment are the terms of an identity, the income saved by consumers being automatically (per double-entry bookkeeping) lent to firms and invested by them (either voluntarily, in the purchase of new investment goods, or forcefully, in the purchase of unsold consumption goods). Hence, saving can neither reduce the amount of available income nor be identified with hoarding, an obsolete concept that has no significant meaning when referred to money income.

What precedes leads us to an unambiguous conclusion: global or macroeconomic demand is defined by the amount of money income generated by production and deposited with the banking system. As a consequence, it is logically impossible to explain the existence of a discrepancy between global demand and global supply by referring to a decrease in consumers' propensity to spend or to any other variation in economic agents' behaviour. Economic crises are a troublesome reality, and deflation is undoubtedly the proof that global demand can fall short of global supply, yet microeconomic analysis can provide no satisfactory explanation of this state of affairs. The problem is of a macroeconomic nature and requires a macroeconomic analysis capable to reconcile the logical identity between global supply and demand with the possibility for global demand to fail to match global supply. This is a far cry from being an easy task, as is confirmed by the following analysis of Say's law.

Say's law and the possibility of crises

Say's law has mainly been rejected on the ground that it does not account for a generalized situation of unemployment. In his *General Theory*, Keynes (1936/1946: 26) argues that Say's law 'is equivalent to the proposition that there is no obstacle to full employment'. It is surprising that Keynes did not realize that the same critique could have been addressed to his own identity. If supply creates its own demand, so the critical argument against Say's law goes, whatever the level of production, the entire output can always be sold, so that nothing prevents firms to increase production until full employment is reached. By the same token, if for any given firm it were enough to produce in order for this firm to sell, and if the identity between Y and $C + I$ did not allow for a numerical difference between its two terms, there would be no room for economic crises to occur.

Let us consider the implications of Say's law on employment first. The law under investigation establishes that every production creates the amount of income necessary and sufficient for the final purchase of

current output. Does this mean that whatever it is produced will be sold on the commodity market? Certainly not. It is true that wage earners are initially credited with an income defining current output. However, nothing forces them or those who take their place as income holders to spend the entire amount of their income in the purchase of their product. The income required for the final purchase of what has been produced is available in the form of bank deposits, but their holders may well decide to save part of it. According to Keynes's principle of effective demand, firms tend to produce the amount of goods and services that are likely to be demanded. Nevertheless, it is always possible that they get it wrong, and that part of what is produced remains unsold. If this is the case, if firms are unable to sell part of their output, they will have to readjust their strategy and modify their production, which may lead to frictional unemployment. Say's law holds good even under these circumstances, as it remains true that production creates the income required for the purchase of the whole output. What happens in this case is that part of current output has to be purchased by firms, which borrow from income holders the amount of money income required to cover the cost of production of their unsold output. The identity between global supply and global demand is preserved and yet firms suffer a loss that could prevent them from increasing the employment level.

This explanation is still far from justifying the existence of crises. Mismanagement might be a cause of frictional unemployment, but cannot explain the insurgence of involuntary or pathological unemployment. We will have to go much deeper into the analysis of our monetary systems to find the origin of crises. For the time being, let us simply show that Say's law as well as the identity between supply and demand at the macroeconomic level (which is but an alternative formulation of Say's law) are compatible with a numerical difference between global supply and demand. A simple consideration should suffice to settle the question: the identity between Y and $C + I$ is substantial and not just nominal. According to this identity, there can be no difference between Y and $C + I$, because each of these two terms defines the other. The result of production is indeed an output defined by a sum of income expressed in wage units. Analogously, demand for consumption and investment goods is defined by the same sum of available income. Hence, both Y and $C + I$ are numerically identical only when expressed in wage units. In an orderly system this would always be the case, and the use of a constant unit of account would grant the respect of the identity both substantially and numerically. Things turn different, however,

when the working of the monetary system does not comply with the logical laws induced by the nature of bank money. In a disorderly system, in fact, the identity still holds good, but it does not respect the arithmetical equality of its two terms.

The easiest way of showing how this may happen is to refer to an inflationary emission of money. Let us suppose that, contravening the laws of money and the logical distinction between money and income, the central bank issues what Schmitt (1975) calls a 'gratuitous' money in order to finance part of the public deficit or as a counterpart of either gold or foreign exchange. A sum of gratuitous or 'empty' money would add up to the income generated by production and would inflate demand, albeit only in nominal terms. Indeed, the amount of income would not be modified by this pathological emission, which would only alter the number of money units that will be used to convey it. The distribution of the same income over an increased number of money units entails a decrease in the purchasing power of each unit, which is exactly what inflation is all about essentially.

A numerical example may prove useful. Suppose production to be measured by 100 wage units. As we know, this means that produced output is defined by 100 wage units and that, in its turn, these 100 wage units define the amount of income available in the system. The identity between global supply and global demand is as follows.

$$Y (100 \text{ wage units}) \equiv C + I (100 \text{ wage units}) \quad (1)$$

If an emission of 10 units of gratuitous or empty money occurs, identity (1) becomes:

$$Y (100 \text{ wage units}) \equiv C + I (110 \text{ money units}) \quad (2)$$

In value terms the identity still holds, for global demand is increased only nominally. The value of 110 money units has remained equal to 100 wage units. This means that an income of 100 wage units is now spread over 110 units of money. The nominal increase in the number of money units is inflationary and defines a decrease in the purchasing power of each money unit.

If 10 units of empty money are added to the 100 units initially associated with produced output, no increase in macroeconomic demand occurs. The income or the purchasing power available remains unaltered, and it is the amount of income available that defines demand. The 10 units of empty money do not make the system any richer. Their impact is on income distribution, not on production.

What matters here is to emphasize the centrality of the identity between global supply and demand. As we have just seen, this identity is compatible with a numerical difference of its two terms, and allows for a precise definition of a pathology such as inflation. This is even more so since, without Say's identity, we would have no means to decide whether an increase in demand is nominal or real, whether an arithmetical difference, like the one of our example, is due to an increase in demand or a decrease in supply. Following Schmitt (1975), let us suppose the payment of wages to be equal to x wage units and the sum spent for the final purchase of current output to be of x' money units. As claimed by the French economist, a numerical difference between x and x' is the necessary condition for a disequilibrium to occur. This is not a sufficient condition, however, because the numerical difference by itself tells us nothing about the economic value of x and x' and their relationship. Say's and Keynes's identities as well as Schmitt's law provide a definition of x and x' as well as their measure in terms of wage units, and establish that their relationship is of a reciprocal nature, where x defines x' and x' defines x . It is because global supply and demand are the two faces of a single reality that x' and x are identical, and that their numerical difference is the mark of an economic disequilibrium.

If we knew nothing about the comparative value of x and x' , nothing will allow us to conclude that the nature of the observed difference between x and x' is that of a disequilibrium. Only the law of the circuit (or Say's law) provides the required piece of information: x is necessarily equal to x' . It results that every positive or negative difference between x and x' causes a disequilibrium, since it is necessarily reabsorbed through a variation in the economic value of the money unit.

(Schmitt 1975: 67, our translation)

The presence or the absence of the identity between Y and $C + I$ is what makes the difference between a logical and rigorous analysis and a confused description of what is perceived to be the real world. In the first case our investigation rests on firm ground and we are guided by a principle whose validity is established by logic; in the second case we have no objective landmark and we are bound to wander from one hypothetical interpretation to another.

Contrary to what Keynes (1936/1946) claimed, the proposition that supply creates its own demand is no remedy against unemployment, but the necessary starting point of any analysis attempting to explain the origin and nature of economic crises. Say's contribution is all-important,

yet it provides no explanation of how crises can arise. His law sets the logical framework of analysis, but it does not tell us how a 'disequilibrium within equilibrium' may appear, and even less how this could be avoided. The problem that has to be solved is not easy: the logical identity between global supply and demand seems to oppose any serious and long-lasting disparity between its two terms, even though it is not incompatible with their numerical difference. The case previously considered in our analysis may be deemed as marginal and largely insufficient to account for the disequilibria characterizing severe economic crises. In most countries, central banks are independent from national governments, and do not finance public deficits through money creation. The direct emission of empty money is no general practice, and it would indeed be wrong to identify it with the cause of recurrent, pathological fluctuations in global demand. How is it possible, then, to work out a theory of crises both consistent with the macroeconomic identity of supply and demand and capable to account for an excess demand and/or an excess supply? Is the traditional attempt to distinguish between supply-side and demand-side economics successful in providing the correct analytical framework required to answer this question? Let us address this problem in the next section.

Supply-side versus demand-side economics: A deceptive alternative

University students are taught that a useful distinction can be drawn between theories that are supply-side oriented and theories emphasizing the role of demand. Roughly speaking, neoclassical theories are said to pertain to the category of supply-side economics, while Keynesian theories are ranged within the realm of demand-side economics. According to this rather far-fetched distinction, economic analysis would be influenced by the choice – mostly ideological – to privilege one over the other of these two approaches. Thus, supply-side economists consider as central the study of whatever situation or event is likely to interfere with free entrepreneurship and prevent production to take place with the less possible hindrance from non-economic factors. Productivity and profit maximization is all that matters in a world that is believed to be potentially the best for entrepreneurs and workers alike. On the other hand, demand-side economists maintain that, according to Keynes's principle of effective demand, consumers' decisions are the determinant factor in the search for the level of output compatible with the equilibrium between Y and $C + I$.

In accordance with this distinction between supply-side and demand-side economics, crises are attributed to disruptions affecting producers' behaviour or consumers' demand and remedies are suggested, which go from reducing public interventions as much as possible to encouraging them as much as possible. Economists either adhere to one of these two dogmatic approaches, or move from one to the other, according to the argument they are dealing with and their personal feelings as to the socio-political implications of either choice. This dismal picture of our discipline is sometimes considered as the unavoidable consequence of an economic reality that is itself very complex and even contradictory. Young economists are thus led to choose between two theoretical frameworks reciprocally exclusive, even though it is widely recognized that none of them is entirely satisfactory and that a synthesis between the two approaches would be preferable to the radical adoption of only one of them. Things would not be too seriously compromised if it were possible to use each single approach to explain a different aspect of economics, that is, if supply-side and demand-side economics were two complementary analyses. Unfortunately, this is not the case, not only because each of them aims at being self-sufficient and all encompassing, but – more important – because none of them can succeed in attaining this objective.

The central feature of supply-side economics is production. Once produced, goods are exchanged in the market and their relative prices are supposed to maximize the utility of economic agents. Goods are exchanged against goods, and the only pre-requisite for their exchange to take place is that they are desirable and/or useful objects. What is missing from this analysis is an explanation of how goods can be sold. Indeed, in this framework money plays no essential role, and is either not distinguished in any fundamental way from a commodity or considered as a simple veil. Income is thus bound to remain totally unexplained. The question that arises quite naturally at this point is how can demand be accounted for in such a theoretical framework. The traditional answer given by neoclassical economists is that in order to demand a good it is necessary to offer another good. The whole of relative price analysis is built around the idea that direct exchange is what determines relative prices, and that the system of general equilibrium allows for a solution because every supply is a demand, and every demand a supply. Known as Walras's law, this principle derives directly from what was already familiar to Aristotle, to wit, that every exchange defines the equivalence of its two terms. The annoying fact, however, is that Walras's law is but a tautology, which applies only when

an exchange effectively takes place. Albeit it is certainly true that when an exchange actually occurs the goods exchanged for one another are equivalent, this equivalence cannot be held to apply during the phase of adjustment that is supposed to lead to the determination of equilibrium prices. Before exchange, prices fluctuate in the search for an equilibrium that is presumed to depend on the adjustment between supply and demand. At this stage Walras's law plays no role, because no actual exchange occurs. As a consequence, it cannot be referred to in order to reduce the number of independent equations making up the system of general equilibrium. The ensuing overdetermination of this system provides a formal proof of its logical inconsistency, that is, of the logical impossibility to determine relative prices through direct exchange.

If, as shown by Schmitt (1984a, 1996, 2012), relative prices are bound to remain undetermined *unless* they are derived from absolute prices, goods cannot become commensurable through relative exchange. Commodities' physical heterogeneity is the first obstacle an economic theory has to overcome. If it fails to do so, it must be abandoned, for it proves to be incapable to provide for a homogeneous measure of economics' proper object of enquiry: currently produced output. Moreover, if goods remain heterogeneous, their supply cannot be given a numerical expression, nor can their demand be known in any objective way. Supply-side economics fails therefore in both its attempts to provide for a general equilibrium theory of exchange, and to establish a logical priority of supply over demand. The reason of its failure is notably the lack of a monetary analysis capable to explain how real goods can be associated to numbers through an absolute exchange that gives numbers a real content, and transforms goods into the object of bank deposits. In the absence of money, goods cannot be made homogeneous and measured by a common (numerical) standard. By the same token, supply and demand remain totally undetermined, and it is of no avail to look for a meaningful explanation of either orderly or disorderly economies.

Demand-side economics avoids the criticism addressed to supply-side economics in that it relies on bank money and defines both supply and demand in monetary terms. Emphasizing the role of demand, the advocates of demand-side economics maintain that consumers' behaviour is all-important in determining the rate of growth of an economy. According to their analyses, a drastic decrease in demand is a major cause of a disorder that, if protracted, inevitably leads to an economic crisis. Conversely, an increase in consumers' demand is held to be a necessary step towards recovery. The relevance given to the marginal propensity to consume within the Keynesian theory of the multiplier is well known,

as is the constantly renewed effort of Keynesian economists to promote economic policies designed to boost aggregate demand.

Is the analysis proposed by demand-side economics correct? The answer to this question requires a brief detour to investigate the role played by economic agents' behaviour in determining global demand. To be sure, economists unanimously agree on the importance of human behaviour for microeconomic analysis. Yet, this is not the point here. What we are concerned with is how macroeconomic demand is established in the first place, and it would be mistaken to maintain from the outset that what is true from a microeconomic analysis of individuals' behaviour point of view also holds at a macroeconomic level. With respect to global demand, the determinant factor we have to consider is income. Whatever the needs and desires of consumers, their demand is bound to remain a virtual magnitude of no impact whatsoever if it is not financed by income. The point that must be clarified is thus whether or not consumers' behaviour can determine the demand effectively resulting from a given amount of income.

Let us suppose available income to be equal to 100 money units. If income holders decide to spend only 80 of the 100 units deposited with the banks, is it correct to infer that global demand is equal to 80 money units? Alternatively said, does saving reduce global demand? The answer is no, for the simple reason that income is saved in the form of bank deposits and bank deposits are always necessarily lent by banks. What matters here is double-entry bookkeeping. Deposits are entered on the liabilities side of banks' ledgers and are balanced by an equivalent entry on the assets side. The meaning of this double entry is straightforward: the income saved by income holders is immediately lent to banks (deposited with them) and lent by them to their clients. When savings are not borrowed by consumers – the only interesting case here – they are lent to firms. In our example, the 20 units of income that are not spent by consumers are borrowed by firms, which use them to cover their deficit formed because of saving. In order to cover their costs of production firms must sell their products. If they are unable to sell part of their current output because of a change in consumers' behaviour, they are forced to purchase it themselves. In the great majority of cases and if the change is confirmed through time, this will simply induce firms to adjust their production in accordance with the consumers' new habits. Yet, what really matters here is not the role of expected or anticipated demand. No one doubts that expectations are an important factor in the determination of firms' future production, but this is of no consequences over current output, which has already been produced. The true

problem concerns the relationship between current supply and demand, and what has to be settled is whether or not an increase in savings reduces current demand.

Keynes's identity between saving (S) and investment (I) guides us towards the correct solution. Current demand is made up by the demand for consumption goods (C) and the demand for investment goods (I). The distinction between these two categories of goods does not rest on their physical characteristics, but on the character of their final purchasers. Savings are spent (invested) by firms, and give thus rise to investment goods, while the income finally spent by consumers defines the amount of current output consisting in consumption goods. An increase in saving is *ipso facto* an increase in the purchase of investment goods owing to the book-entry definition of income in the form of bank deposits (see Chapter 2). The decrease in the demand for consumption goods is therefore immediately compensated by an equivalent increase in the demand for investment goods. Hence, we are led to conclude that a variation in consumption has no effect on total demand, whose amount is unequivocally determined by $C + I$.

Quantum macroeconomics confirms the conclusion derived from Keynes's identity. Current supply is but the result of current production and is from the outset identified with the amount of wages paid to workers. Because of double-entry bookkeeping, and in line with the flow nature of money, income is formed as a bank deposit and is immediately lent to firms. At the very moment wages are formed, they are thereby invested by firms in the purchase of current output. Current supply is therefore matched at once by an equivalent demand exerted by firms. The whole of produced output is transformed into a stock, which, resulting from the investment of firms, may be considered as made up of investment goods. In this respect it can even be claimed that at the beginning output consists of investment goods only. Inasmuch as income holders spend what was initially formed as a bank deposit, the stock is reduced and investment goods are definitively transformed into consumption goods. Hence, it is clear that whatever its composition, the totality of current output is demanded by consumers and/or by firms.

Finally, let us observe that it is inconsistent to maintain that saving can at the same time reduce global demand and finance investment. Since, in compliance with the principle of double-entry bookkeeping, the part of current income that is not spent for the purchase of consumption goods is necessarily deposited with banks and thus lent to firms, what is not spent by consumers is spent by firms, to wit, it is invested. It hence follows that macroeconomic saving can never differ

from macroeconomic investment. This result is perfectly in line with the position adopted by Keynes following his debate with Robertson after the publication of his *Treatise on Money* and during the preparation of *The General Theory*: 'for the community as a whole, investment and saving are necessarily, and by definition, equal' (Keynes 1973a: 476). Quantum monetary macroeconomics confirms Keynes's intuition and strengthens it by showing that the identity between macroeconomic saving and investment is not simply a matter of nominal definition but the unavoidable result of double-entry bookkeeping applied to money and income. As global demand is measured by the amount of income that finances it, and since saving finances investment, it appears that saving can never cause a pathological shortage of global demand.

Beyond the distinction between demand-side and supply-side economics

Classical economists and Keynes on one side, and neoclassical economists on the other side, may be viewed as the advocates of two conflicting and reciprocally exclusive approaches, none of which is entirely satisfactory. Alternatively, they can be considered as the exponents of two necessary phases in a dialectical process converging to a synthesis that encompasses their analyses and that, by resolving the contradiction between the two approaches, goes beyond them. What we are referring to here is not the neoclassical endeavour to assimilate Keynesianism first to the neoclassical and then to the new neoclassical synthesis. Both these attempts consist in imposing general equilibrium as a theoretical framework of analysis and in reducing macroeconomics to an aggregation of microeconomic variables. Neither the neoclassical nor the new neoclassical synthesis goes beyond neoclassical analysis, their attempt being that of incorporating some Keynesian and New Keynesian features in models whose foundation remains that of real business cycle (RBC) models. 'The New Neoclassical Synthesis is defined by two central elements. Building on new classical macroeconomics and RBC analysis, it incorporates intertemporal optimization and rational expectations into dynamic macroeconomic models. Building on New Keynesian economics, it incorporates imperfect competition and costly price adjustment' (Goodfriend and King 1997: 255). The models worked out by New Keynesian economists in their attempt to provide a New Keynesian synthesis are substantially the same as those of the new neoclassical synthesis: they are based on a dynamic stochastic general equilibrium (DSGE) framework and on the structure of RBC models. 'The New Keynesian modelling approach [...] combines the DSGE structure

characteristic of RBC models with assumptions that depart from those found in classical monetary models' (Galí 2008: 5).

The distinction between demand-side and supply-side economics is a direct consequence of that between money and output. By emphasizing the role played by money in the determination of demand, advocates of demand-side economics maintain that production adjusts to the expected level of effective demand, whereas proponents of supply-side economics claim that production is the leading factor, and that monetary disturbances such as inflation or government decisions to increase marginal tax rates hamper economic growth. Generally speaking, demand is expressed in monetary terms and is determined by factors substantially different from those determining the amount of economic output.

The dialectical overcoming of the distinction between demand-side and supply-side economics is made possible by the discovery of the mutual dependence characterizing money and output. In a monetary economy of production, money is literally created only when it is 'coupled' with current output. Isolated from output, money has no *raison d'être*. Reciprocally, output exists as an economic product only insofar as it is associated to money. As soon as it leaves its monetary form, output ceases to define an economic product and continues its existence as a physical object. As an economic product, output has or is a positive exchange-value; as a physical object it is or has a use-value. It is at the very moment output acquires a monetary form that it starts existing as an economic entity and that the concepts of supply and demand become meaningful in the market for produced goods and services. The payment of wages is the event enabling the structural coupling of money and output, that is to say, allowing for an economic production to occur. In economics, production is an instantaneous event, that is, a payment, which, by giving money a real content and output a numerical form, brings currently produced output into existence.

The crucial point here is that money and output are the two aspects of one and the same reality, and that the existence of an object of investigation proper to economics is possible only because money and output are so closely entangled as to define the terms of an identity. Putting the logical identity between money and output to the fore, quantum macroeconomics gives a new life to Keynes's identity between Y and $C + I$, and establishes a relationship of reciprocal causality between global supply and demand. If national income, Y , is always and necessarily equal to the income finally spent for the purchase of consumption and investment goods, $C + I$, it follows necessarily that global demand

and supply define one another. This amounts to saying that global supply is the cause of global demand no more and no less than global demand is the cause of global supply. Hence, no logical priority can be attributed either to supply or demand, which makes the distinction between supply-side and demand-side economics obsolete.

The synthesis epitomized by Schmitt's quantum macroeconomic analysis encompasses the classical and the neoclassical points of view, and vindicates Keynes's intuition about the identity between the payment of the macroeconomic costs of production and the final purchase of output. In a given period, global supply is defined by the value of current national output, whose measure is determined by its macroeconomic costs of production. Now, Schmitt (1984a) proves what was already intuited by the Classics as well as by Keynes, namely that human labour is the only macroeconomic factor of production. Thus, the payment of wages is the only macroeconomic cost of production and, by way of inference, wages are the macroeconomic measure of national output. Using Keynes's notation, we would write $Y = C' + I'$, where $C' + I'$ stands for the amount of wages paid for the production of consumption and investment goods. Keynes's initial identity takes now the form of $C + I = C' + I'$ and establishes the necessary equality of the income created by production and the income spent for the final purchase of consumption and investment goods.

Being the terms of an identity, macroeconomic supply, Y or $C' + I'$, and macroeconomic demand, $C + I$, can never differ from one another when they refer to the same output. Indeed, the measure of macroeconomic supply is at the same time the measure of macroeconomic demand. Every economic analysis whose aim is to explain the orderly and the pathological working of our economies has to start from here. A numerical difference between global supply and demand can result from the unilateral variation of one side or the other of the equation defining the relationship between these two magnitudes. For example, an excess demand can arise from an autonomous variation on the demand side, yet this variation would be meaningless, if it were not referred also to the supply side and if the two sides did not represent the two terms of an identity. Likewise, an excess supply is symptomatic of a well-defined pathology, only if it is interpreted with reference to the underlying, substantial identity between supply and demand at the macroeconomic level.

Quantum macroeconomics provides a new framework of analysis based on the macroeconomic identity between supply and demand. The traditional distinction between supply-side and demand-side economics

appears to be far-fetched and is superseded by a synthetical view where both sides are equally relevant. The aim of the quantum macroeconomic synthesis is to show that neither the neoclassical dichotomy between real and monetary variables nor the orthodox interpretation of Keynes in terms of general equilibrium provides the analytical tools required for a correct understanding of our monetary economies and of the pathologies afflicting them. What traditional theories are lacking is a correct understanding of money and its relationship with output. This is why they emphasize either the supply or the demand side in their attempt to reproduce through modelling the working of our economic systems. Despite their reciprocal interrelations, money and output are traditionally conceived of as autonomous entities determined by two different sets of institutions: monetary authorities and banks on one side, and firms on the other side. In reality, this theoretical representation is at odds with facts. Indeed, factual observation shows that money, issued by banks in compliance with the principle of double-entry bookkeeping, exists only to the extent that it is merged with current output. The identity between money and output and, consequently, between macroeconomic supply and demand, is rooted in the very nature of our monetary economies of production. It is therefore necessary to go beyond the old-fashioned distinction between supply-side and demand-side economics, to develop a unified theory based on the macroeconomic identity of supply and demand.

Part III

The Monetary Macroeconomics of Crises

This page intentionally left blank

7

Capital Accumulation and Economic Crises

In the first section of this chapter we will examine how some of the most renowned economists of the past have explained capital. Undoubtedly, capital is one of the central concepts of economics and yet there is still no consensus among economists on how to define it. The questions raised by Smith (1776/1991) concerning the logical relationship between capital and saving, and between circulating and fixed capital, are no longer on the agenda, and yet they have never been satisfactorily answered. This is even more so with regard to the relationship between capital and economic value. Well-known economists of the past such as Ricardo, Marx, Walras, Böhm-Bawerk, and Keynes have addressed this question and, with the notable exception of Walras, have reached the conclusion that capital cannot be considered as a direct source of economic value. However, none of them denies the impact capital has on prices, and it is since Ricardo's (1817/1951) *Principles* that economists are aware that a satisfactory theory of value must account for the presence of capital. In particular, both Ricardo and Böhm-Bawerk (1889/1959) emphasize the role played by time in enabling capital to be an indirect source of economic value. Keynes's analysis of capital is another important contribution to a correct understanding of this concept and encapsulates all the deepest insights of his predecessors concerning the role of saving and time. By introducing these elements into a theoretical framework where the presence of money and banks is essential, Keynes opens the way to the modern macroeconomic analysis of capital and to Schmitt's (1984a) quantum macroeconomic approach.

The second and third sections of this chapter are devoted to the contribution of Schmitt's quantum analysis to the understanding of capital and of its role in the genesis of economic crises. In these sections, therefore, we will provide a factual and conceptual explanation of how

capital is formed in a monetary economy of production. In particular, we look at the reason why fixed capital formation is at the origin of a pathology leading to inflation and unemployment. It is through the investment of profit that fixed capital is formed, and quantum analysis shows that, to date, the systems of domestic payments have been suffering from a structural anomaly that introduces a structural disorder into the economic system. This disorder is the cause of both inflation and unemployment, which appear to be the twin outcomes of the same pathology. Leaving aside any microeconomic consideration emphasizing economic agents' behaviour, quantum analysis provides a macroeconomic theory of capital in line with the main findings of monetary macroeconomics and explains the outbreak of economic crises as ensuing from the disharmony between the present system of payments and the logical laws deriving from the peculiar nature of bank money.

Lessons from the history of economic thought

The concept of capital is one of the most important and controversial in the history of economic thought. This section offers some considerations based on the contributions of a small number of authors whose analysis of capital has contributed to the evolution of this concept. Our aim is to stress the relevance of these contributions for the correct understanding of capital accumulation and of its impact on economic growth.

Turgot and Smith

According to Schumpeter (1954/1994: 323), even though Quesnay (1758/2005) emphasized the role played by Nature, he 'may be credited with laying the foundations of a capital theory'. In Quesnay's *Tableau économique*, advances are at the origin of a circular process of production whose surplus, generated by the intervention of Nature, enables their reproduction. 'The advances are goods – to live on or produce with – though their quantity may be expressed in terms of money, and they are precisely what capital means in one of the many sense of the word' (ibid.: 323). As pointed out by Schumpeter, it is to Turgot (1767/2011) that we owe the first fruitful insights into the theory of capital. Indeed, he was the first to clearly state that saving is the necessary prerequisite for any process of capital accumulation. In his *Observations on a Paper by Saint-Péray*, Turgot (1767/2011) went as far as to claim that savings are immediately spent and converted into capital,

thus introducing the idea – later taken up by Keynes and confirmed by quantum macroeconomics – of the logical identity between saving and investment.

Turgot's contribution was taken over by Smith (1776/1991: 301), who, at the beginning of his analysis on capital accumulation, argues that '[a]s the capital of an individual can be increased only by what he saves from his annual revenue or his annual gains, so the capital of a society, which is the same with that of all the individuals who compose it, can be increased only in the same manner'. Smith develops at length the idea that parsimony is at the origin of capital, arguing that saving leads to the formation of the stock of goods required for the maintenance of workers producing capital goods. 'Parsimony, by increasing the fund which is destined for the maintenance of productive hands, tends to increase the number of those hands whose labour adds to the value of the subject upon which it is bestowed' (ibid.: 301). The message is clear: in order to produce capital goods it is first necessary to create a fund, and this can be done only through saving, that is, by renouncing to consume the totality of currently produced output. Before capital goods are produced, output consists in consumption goods only. On these conditions, the fund created through saving is but a stock of consumption goods. In the following period, workers can partly devote their activity to the production of capital goods, their real remuneration being guaranteed by the stock formed thanks to society's parsimony.

Smith expresses Turgot's intuition about the immediate conversion of savings into capital as follows. 'What is annually saved is as regularly consumed as what is annually spent, and nearly in the same time too; but it is consumed by a different set of people' (ibid.: 302). Smith seems to distinguish between the conversion of savings into capital, which is immediate – 'that portion which he [a rich man] annually saves [...] is immediately employed as capital' (ibid.: 302) – and the consumption of the goods stocked into the fund created through saving, which may take place slightly later or 'nearly in the same time' (ibid.: 302).

In order to understand Smith's analysis better, let us distinguish between a logical or analytical sequence of events and their chronological sequence. Starting from a production of consumption goods, the first logical step required to provide the economy with a positive amount of capital goods is saving. The consequence of saving is the formation of a stock of consumption goods. This stock is then used for the maintenance of workers producing capital goods. Although these logically distinct events may take place at the same time, it is irrelevant whether the consumption goods stocked into the fund created by

saving are purchased immediately or not by workers employed in the production of capital goods. The crucial points of Smith's analysis are that:

1. saving is a necessary step towards capital accumulation;
2. saving is immediately converted into capital;
3. the production of capital goods is likely to take place at a later moment in chronological time.

What needs to be clarified is how to reconcile points 2 and 3, namely how it is possible to maintain that capital is formed at the very moment part of current income is saved and at the same time that capital goods may not have been produced yet? The answer to this question lies in Smith's distinction between circulating and fixed capital.

Let us make it clear from the outset: Smith's definition of circulating capital is very broad and, as such, of little help. However, there is a specific meaning of the concept of circulating capital that, if retained, gives a new life to Smith's analysis of capital. In *The Wealth of Nations*, Smith (1776/1991) starts his chapter on the division of stock by defining circulating capital as that set of goods that yield revenue to their owners only if these goods are exchanged, that is, if they circulate. 'The third and last of the three portions into which the general stock of the society naturally divides itself, is the circulating capital; of which the characteristic is, that it affords a revenue only by circulating or changing masters' (ibid.: 247). Smith then specifies the different kind of goods that he considers as part of circulating capital. 'The circulating capital consists in this manner, of the provisions, materials, and finished work of all kinds that are in the hands of their respective dealers, and of the money that is necessary for circulating and distributing them to those who are finally to use or to consume them' (ibid.: 247). If we leave money aside, the goods entering the category of circulating capital are but the set of consumption goods currently produced by a given economy. By contrast, fixed capital is defined as that set of goods that 'affords a revenue or profit without circulating or changing masters' (ibid.: 246). The crucial point here is that the production of consumption goods (circulating capital) is logically prior to the production of fixed capital goods. If circulating capital is thus defined as the totality of consumption goods and if saving is still considered as the necessary step to form a positive capital, we are led to the conclusion that current income resulting from the production of consumption goods is saved as soon as it is formed and that its corresponding output is immediately stocked into a fund called

circulating capital. As we shall see in the next section, this first and broad concept of circulating capital acquires its full significance when it is referred to the flowing of time, and is the necessary consequence of the implementation of double-entry bookkeeping. For the time being, let us follow Smith in his investigation of the way circulating capital can be related to fixed capital.

As Smith tells us, circulating capital goods are either purchased by income holders (consumers) or placed into fixed capital. '[P]rovisions, materials, and finished work – are, either annually, or in a longer or shorter period, regularly withdrawn from it, and placed either in the fixed capital or in the stock reserved for immediate consumption' (ibid.: 248). What is the meaning of the word 'placed' used by Smith? If the meaning were 'transferred from one category of goods to another', the production of circulating capital goods would be inclusive of that of fixed capital goods, and Smith's distinction between circulating and fixed capital would lose most of its significance. A more fruitful interpretation consists in saying that part of circulating capital goods can be used in order to support the production of new capital goods. Such an interpretation is in line with the claim originally made by Turgot (1767/2011) that capital results from saving, and with a definition of circulating capital that identifies it with a stock of saved-up consumption goods. This is the definition proposed by Turgot of what he called 'moveable wealth': 'Possessions of this kind, resulting from the accumulation of annual produce not consumed, are known by the name of *moveable wealth*' (Turgot 1767/2011: 30). This is also what is implied in Smith's claim that '[e]very fixed capital is both originally derived from, and requires to be continually supported by a circulating capital. All useful machines and instruments of trade are originally derived from a circulating capital, which furnishes the materials of which they are made, and the maintenance of the workmen who make them' (Smith 1776/1991: 248).

The idea conveyed by Smith's quotation is clear: the production of fixed capital goods requires the existence of a stock of saved-up consumption goods (circulating capital), which will be purchased by workers producing instrumental goods. Thanks to the distinction between circulating and fixed capital it is now possible to reconcile the claim that saving is at once transformed into capital and the claim that the production of capital goods may take place later. When firms stock part of the currently produced consumption goods, these goods are immediately transformed into circulating capital, whether or not the production of fixed capital goods is already on the way. The consumption goods

stocked by firms will then be used in ‘the maintenance of the workmen who make them [fixed capital goods]’ (ibid.: 248). This description is both logical and chronological, yet the latter is merely didactical, and can be dropped without causing any harm to Smith’s analysis. In fact, the production of fixed capital goods can take place simultaneously with that of consumption goods, and circulating capital may immediately be replaced by fixed capital. Whether we choose to analytically separate the formation of circulating capital from the production of fixed capital goods or not, the result is unchanged: ‘[e]very fixed capital is [...] originally derived from [...] a circulating capital’ (ibid.: 248).

Ricardo and Marx

In Chapter 5 of his *Principles*, Ricardo (1817/1951: 95) defines capital as ‘that part of the wealth of a country which is employed in production, and consists of food, clothing, tools, raw materials, machinery, &c. necessary to give effect to labour’. From this definition it appears that Ricardo is not particularly interested in classifying capital goods under the heads of circulating or fixed capital. All the goods entering, directly or indirectly, into the production of final output are capital goods, and Ricardo’s main concern lies with the impact that the use of capital goods has on the value of output.

As a classical economist, Ricardo endorsed the idea that value is determined by labour. ‘I am fully persuaded that in fixing on the quantity of labour realised in commodities as the rule which governs their relative value we are in the right course’ (Ricardo 1953: 344). Yet, the Anglo-Portuguese economist was well aware that commodities’ value cannot be determined by relating only to the quantity of labour directly incorporated in them. Capital goods must also be taken into account. Ricardo’s first step is to include part of the labour required for the production of capital goods into the value of output: ‘the exchangeable value of the commodities produced would be in proportion to the labour bestowed on their production; not on their immediate production only, but on all those implements or machines required to give effect to the particular labour to which they were applied’ (Ricardo 1817/1951: 24).

This is not enough, however, since capital goods are of different durability and have different rates of turnover. As claimed by Ricardo,

[t]he wheat bought by a farmer to sow is comparatively a fixed capital to the wheat purchased by a baker to make into loaves. One leaves it in the ground, and can obtain no return for a year; the other can get it ground into flour, sell it as bread to his customers, and have his

capital free to renew the same, or commence any other employment in a week.

(*ibid.*: 31)

How are we to account for all these different kinds of capital goods, and how are we to explain their impact on the value of the commodities to whose production they contribute? Ricardo, 'by what almost amounted to a flash of genius' (Schumpeter 1954/1994: 636), saw that their common denominator was time, more precisely, that their economic difference could be reduced to the period of time elapsing between their production and their reappearance in the form of final output.

As a matter of fact, Ricardo put forward two different problems: one related to the different proportions between labour and capital invested in the production of commodities, and another one derived from the different periods of time required to bring to the market different commodities of equal value. Ricardo solved the first problem by assuming that the value of capital goods is transferred to final output to the extent that they are used up in its production. 'Here then are capitalists employing precisely the same quantity of labour annually on the production of their commodities, and yet the goods they produce differ in value on account of the different quantities of fixed capital, or accumulated labour, employed by each respectively' (Ricardo 1817/1951: 34). The second problem is the one that induced Ricardo to consider time beyond the then traditional concept of labour-time. The following quotation shows how by the different degrees of durability of capital he means the difference in the periods of time during which capital must be invested, and not the difference in the ratio of labour and fixed capital.

On account then of the different degrees of durability of their capitals, or, which is the same thing, on account of the time which must elapse before one set of commodities can be brought to market, they will be valuable, not exactly in proportion to the quantity of labour bestowed on them, – they will not be as two to one, but something more, to compensate for the greater length of time which must elapse before the most valuable can be brought to market.

(*ibid.*: 34)

Ricardo's idea was that the value of commodities produced by labour with the support of instrumental goods cannot be explained by the quantity of labour-time corresponding to the activity of workers giving

a new utility form to matter and energy and to the activity of workers producing the raw materials and that part of fixed capital goods used up in the production of final output. The period of time during which capital has to be invested before firms can recover it in the sale of final output has also to be taken into account. This is what Ricardo tells us in one of his most quoted passages:

I sometimes think that if I were to write the chapter on value again which is in my book, I should acknowledge that the relative value of commodities was regulated by two causes instead of by one, namely, by the relative quantity of labour necessary to produce the commodities in question, and by the rate of profit for the time that the capital remained dormant, and until the commodities were brought to market.

(Ricardo 1953: 194)

Ricardo's analysis of capital brings to the fore the role played by time, and is confronted with the challenge to reconcile the Classics' claim that labour is the unique macroeconomic factor of production with the necessity to account for the increase in value owing to the presence of circulating and fixed capital goods. His solution consists in assuming that profits are proportionally distributed among firms on the basis of their total investment in labour and capital. Since, according to Ricardo and the Classics, circulating capital and fixed capital goods are but accumulated labour, the value added by capital on top of labour is due to the need to remunerate firms for their investment. Unfortunately, Ricardo was not up to the task he set to himself, and did not provide an entirely satisfactory solution to the problem of value. Yet, his intuition proved very fruitful, as is confirmed by the analysis of capital advocated by Böhm-Bawerk (1889/1959) and Schmitt (1984a). Before considering Böhm-Bawerk's contribution, let us briefly comment on the work of one of Ricardo's greatest admirers: Marx.

Ricardo's conundrum about the role played by capital finds its most worrisome version in Marx, who replaced the classical distinction between circulating and fixed capital with that between variable and constant capital, and who considered workers' labour-power as the unique source of value. True to what he believed to be his main contribution to the analysis of capitalism, Marx was confronted with the task to incorporate constant capital into his labour theory of value. The problem, analogous to the one addressed by Ricardo, is as follows. Measured by labour time, value is determined by the new activity of

workers and by that part of labour accumulated in the form of constant capital that is transferred to final output. Yet, the capital invested by firms in the payment of workers is substantially different from that invested in the purchase of fixed and circulating capital goods. The former, which Marx calls variable capital, gives firms the right to exploit workers' labour-power, whereas the latter, called constant capital, gives them the ownership over previously produced instrumental and intermediate goods. Now, the purchase of labour-power enables firms to derive a surplus value from their investment in variable capital, whereas the investment in instrumental and intermediate goods can only lead to a transfer of value. This transfer is equivalent to the incorporation in final output of part of the value of constant capital, that is, of that part of 'dead' labour used up in the production of final goods and services. Through their labour, workers are at the origin of value. Yet, when firms pay workers their due (wages), they acquire the labour-power of the latter, which enables firms to exploit labour-power beyond what is strictly required to reproduce it: surplus-value is precisely what firms obtain for free when workers' working time exceeds that necessary to reproduce their labour-power. As labour alone is the source of value, it appears that, if goods were to be exchanged according to the strict labour theory of value, firms investing more in variable than in constant capital would realize a much higher profit than those investing more in constant than in variable capital.

What Marx calls the organic composition of capital is the proportion between variable and constant capital. 'By the composition of capital we mean, as already stated in Volume 1, the ratio between its active and its passive component, between variable and constant capital' (Marx 1894/1981: 244). Having centred all his analysis on the concepts of labour-power and surplus-value, Marx had to come to terms with the indisputable fact that profit is proportional to the capital invested, irrespective of its organic composition: 'the whole difficulty arises from the fact that commodities are not exchanged simply as commodities, but as the products of capitals, which claim shares in the total mass of surplus-value according to their size, equal shares for equal size' (ibid.: 275). As Marx himself put it, the labour theory of value, as exposed in Book 1 of *Capital*, 'appears incompatible with the actual phenomena of production' (ibid.: 252), since it would hold good 'only for capitals of the same organic composition' (ibid.: 252). Moreover, Marx, following Ricardo, observes that a correct integration of capital into the theory of value must also take into account differences in capital turnovers. Capitals of the same organic composition may differ as to the period

of time required for their reconstitution, which means that an accurate determination of the average rate of profit has to incorporate what Böhm-Bawerk (1889/1959) was later to call the roundabout methods of production.

Besides the differing organic composition of capital, i.e. besides the different masses of labour, and therefore, other things being equal, of surplus labour as well, set in motion by capitals of the same size in different spheres of production, there is a further source of inequality between rates of profit: the variation in the length of capital turnover in the different spheres of production.

(Marx 1894/1981: 250)

Forced to rethink his initial labour theory of value, Marx devoted his efforts to the search of a solution that would avoid giving it up. Still convinced of its substantial soundness, he attempted to reconcile it with the '[f]ormation of a general rate of profit (average rate of profit)' (ibid.: 254) through the 'transformation of commodity values into prices of production' (ibid.: 254). If commodities cannot be exchanged according to their labour values, Marx claimed, their prices do nevertheless respect the priority of values, which remain the pivotal element of their determination. Profits are distributed according to the general rate of profit, yet their amount is determined by that of surplus-value, so that, for the economy as a whole, the equivalence between values and prices still holds good.

Unfortunately, Marx's attempt to safeguard the centrality of the classical labour theory of value did not prove successful. The problem of the transformation of values in prices has remained with Marxian economists since Marx's death and is bound to remain unresolved, because it rests on a wrong analysis of profit and value. The identification of profit with surplus-value is Marx's biggest mistake. The assumption that capitalism is characterized by the existence of a commodity called labour-power is totally alien to the real world of economics. Moreover, the definition of value as a quantity of labour-power associated with the idea that firms' profit is formed as the difference between total labour time and the labour time required for the reproduction of workers' labour-power encounters two insurmountable obstacles, namely the heterogeneity of labour and the impossibility to explain the 'monetary realization' of surplus-value.

Despite this unrewarding result, Marx's analysis of capital remains a stimulating contribution and provides fruitful insights, such as the law

of the 'tendential' fall in the rate of profit, which are still central for our understanding of economic crises. But let us advance a step further and see what have been the main contributions of Walras, Böhm-Bawerk, and Keynes to the theory of capital.

Walras, Böhm-Bawerk, and Keynes

The founding father of general equilibrium analysis rejected the classical definition of value in terms of labour time and replaced the concept of absolute value with that of relative prices. As for the problem of capital, Walras (1874/1984: 267) considers '[t]he determination of the prices of capital goods [as] the third major problem of the mathematical theory of social wealth'. Having first analysed exchange and then production, Walras applies the same method to capital formation: the search for equilibrium through a set of simultaneous equations. He starts by introducing a '*capital goods market*, where capital goods are bought and sold' (ibid.: 267), and by claiming that '[t]he price of a capital good depends essentially on the price of its services' (ibid.: 267). Then, after observing that capital goods yield an income, he explains that the determination of prices of different capital goods yielding the same income must account for their amortization and for an insurance premium enabling us to consider all capital goods as rigorously identical with respect to wear and tear as well as to accidental loss. Reduced to its essential elements, the problem of equilibrium prices of new capital goods requires, according to Walras, the simultaneous solution of a set of equations concerning (1) the equality between the value of the new capital goods and the value of saving, and (2) the equality between their prices and their cost of production. The first condition is derived by Walras from the necessity to determine 'the rate of net income and consequently for the determination of the prices of capital goods' (ibid.: 269). The second condition is the necessary consequence of the fact that 'new capital goods are products' (ibid.: 269), so that, like any other product, their price determination must respect what Walras calls the equilibrium in production, that is, 'a state in which the selling prices of products equal the costs of the productive services that enter into them' (ibid.: 224).

Apart from the claim that 'equilibrium in capital formation [...] will be established *effectively* by the reciprocal exchange between savings to be accumulated and new capital goods to be supplied' (ibid.: 282–3), which supports the idea that capital is derived from saving, nothing relevant is said by Walras with respect to the nature of capital. He assumes, *a priori*, that capital is a factor of production as are labour and land, and that, as such, it is at the origin of an income (interest),

which at equilibrium defines its price. All in all, Walras's analysis of capital does not provide any fruitful insight and could easily be discarded were it not for the tremendous impact it had and still has on mainstream economics. Its great appeal lies with the mathematical formalization provided by Walras's approach, and the belief that economics can derive its rigour only from the use of mathematical and statistical modelling. Blinded by the technical sophistication of their models, Walras's followers do not see the poverty of their conceptual analysis and underestimate the relevance of the critiques addressed to it. In this respect, let us simply recall the accusation of circularity (Garegnani 1972), which cannot be rejected simply on the ground that all the variables of the system are determined simultaneously (an acceptable solution in mathematics, but sheer nonsense in economics). Last but not least, Walras's decision to exclude money from his analysis of capital shows very clearly how far from reality his theory is. His claim that 'the key to the whole theory of capital is to be found in thus eliminating capital loans *in the form of numéraire* so that attention is directed exclusively to the lending of capital *in kind*' (Walras 1874/1984: 290) leaves no room for doubt. Money and monetary capital are considered by Walras as 'nothing but a superfoetation in theory' (ibid.: 290) 'however useful in practice' (ibid.: 290). A striking example of the dichotomy between theory and reality introduced into economics by general equilibrium analysis, this sentence is symptomatic of the difficulties encountered by Walras. Was Böhm-Bawerk able to avoid them?

This question is relevant for at least two good reasons: (1) Böhm-Bawerk is one of the greatest exponents of the Austrian School; (2) he is also considered as 'one of the great architects of economic science' (Schumpeter 1954/1994: 847) and a specialist in the topics of capital and interest.

With respect to capital, Böhm-Bawerk's contribution may be seen as an attempt to solve Ricardo's and Marx's conundrum about the impact of capital on the value of final output by referring to the concept of utility. The Austrian economist was both influenced by and very critical towards the analysis advocated by classical economists. In particular, he did not embrace the Classics' labour theory of value and, following Menger (1950), he claimed that commodities' exchange-value is determined by their (marginal) utility. Yet, unlike Walras (1984), he shared the Classics' idea that capital is not a macroeconomic factor of production, because it is not a direct source of value. 'Capital does not independently deliver an impulse, it merely transmits an impulse delivered by originary productive forces [labour and natural agents]'

(Böhm-Bawerk 1889/1959: 95). As observed by Schumpeter (1954/1994: 901), Böhm-Bawerk 'fought the idea that "physical" capital is a distinct factor of production, capable of being treated on the same plane with the "original" factors, labor and natural agents'. As a result, Böhm-Bawerk was faced with the challenge to explain how the determination of value can account for the existence of different capitals employed during variable lengths of time in the production of final output without considering capital as a factor of production as labour (and natural resources). The problem is essentially the same as the one addressed by Ricardo and Marx, yet Böhm-Bawerk rejects Marx's distinction between variable and constant capital as well as the Classics' definition of value in terms of labour time.

Böhm-Bawerk's starting point was, like the Classics', the formation of a wage-fund made up of consumption goods. Necessary to provide the means of subsistence required for the production of fixed capital goods, this initial stock of consumption goods is made possible through saving and represents 'a subsistence fund which makes circuitous production possible' (Böhm-Bawerk 1909/1959: 73). This stock, which Böhm-Bawerk defines as a reserve of productive power, is employed in the production of instrumental goods and, together with the employment of labour and natural resources, enters into the production of final output. According to Böhm-Bawerk, both fixed and circulating capital goods are but 'an aggregate of intermediate products' (ibid.: 73). Now, the investment of labour as well as of circulating and fixed capital (as intermediate products) during variable intervals of time gives rise to what Böhm-Bawerk calls the roundabout process of production. The concept of roundabout methods of production is what enables Böhm-Bawerk to maintain that, although it is not itself a source of value, fixed capital brings about an increase in commodities' exchange-value, because its use has an impact on their production period. In his own words, capital is 'indirectly productive because it makes possible the adoption of new and fruitful roundabout methods of production' (Böhm-Bawerk 1889/1959: 101). Time plays a central role here, and Böhm-Bawerk is rightly considered as one of the authors who have most correctly emphasized the close correlation that exists between capital and time. According to him, value is influenced by the average period of production, that is, by the average period during which capital is invested in a given production. Böhm-Bawerk was not interested in the increase in use-values that the use of circulating and fixed capital goods brings about. He was perfectly aware of the fact that the use of instrumental goods multiplies the number of finished goods produced during

a given interval of time, which amounts to saying that each single finished good is produced in a shorter period of time. The increase in value Böhm-Bawerk refers to in his analysis is a macroeconomic increase in the exchange-value of the total of produced output. More complex roundabout methods of production lead to an overall increase in value and to the production of a larger number of finished goods.

If Böhm-Bawerk's analysis were to end here, it would not substantially differ from Ricardo's. His roundabout period of production resembles Ricardo's period of time during which different capitals are invested before commodities are brought to market, so that it would seem 'possible to claim that period of production and labour embodied in capital goods are two ways of accounting for or describing the same method of measurement of capital' (Garegnani 1972: 28, our translation). In reality, Böhm-Bawerk does not endorse the Classics' analysis of capital, nor does he agree with their measurement of value, included that of capital goods, in terms of labour time. His solution consists instead in evaluating goods on the basis of their utility, and in claiming that the roundabout methods of production are likely to determine an increase in value because of individuals' propensity to prefer present over future goods. Böhm-Bawerk (1889/1959: 268) claims indeed that '[w]e systematically undervalue our future wants and also the means which serve to satisfy them'. This means that a premium is formed, in favour of present goods, which is at the origin of a discrepancy between the expected value of future goods and their effective value. According to Böhm-Bawerk, this discrepancy varies in accordance with the duration of the roundabout period of production and explains the increase in value due to the use of capital and the passing of time.

Böhm-Bawerk explanation is not really satisfactory, because it rests on a psychological principle that can, at most, account for a generalized misconception of value. His attempt to show that capital is indirectly productive since it involves 'the adoption of new and fruitful roundabout methods of production' (ibid.: 101) fails, because the impact of roundabout periods of production over value is merely reduced to a systematic undervaluation instead of resulting in an objective variation of value itself. If it is true that economic value cannot be identified with any physical dimension, labour-time included, it is also true that it cannot be transformed into a psychological dimension. Once again, what is still missing is a rigorous analysis of money and of its relevance for a correct macroeconomic analysis of value, production, and capital.

Despite these major shortcomings, Böhm-Bawerk's analysis of capital remains a cornerstone in the history of economic thought, both for the

centrality attributed to the concept of time, and for the uncompromising refusal to consider the physical productivity of capital as a source of value.

Let us conclude our overview with the contribution of Keynes.

In the first draft of Chapter 1 of his *Treatise on Money*, Keynes (1973a: 19) claims that '[a] supply of *new* capital, whether in the form of finished goods or of goods in process, can only come into existence in so far as those who have claims on the community's flow of income are willing to *defer* their claims, i.e. out of "savings"'. His position is thus perfectly in line with that of most authors of the past, who very quickly fathomed that saving is the first necessary step in the process of capital formation. Yet, Keynes's analysis emphasizes the role played by banks from the outset: 'the fluctuating demands of industry for claims on current real income, which can, by furnishing real wages to labour and in other ways, be converted into goods in process, are mainly satisfied through the banks' (ibid.: 19). Keynes's contribution consists in having clearly perceived the centrality of banks as intermediaries between firms and income holders, and of bank money as the vehicle through which saved-up income is conveyed to firms and invested in the production of capital goods.

A correct understanding of capital requires its analysis to be carried out both in real and financial terms. The constitution of a real wage-fund is not enough to explain the formation of capital; a further step is needed, which enables income to be saved in a period and spent in another, that is, to establish a financial bridge between present and future. One of Keynes's great intuitions is that the existence of capital is closely related to that of banks, that without banks this 'bridge' could not exist and no capital would be formed. His argument is clear:

1. capital 'can only come into existence [...] out of "savings"' (ibid.: 19);
2. saving consists in deferring 'claims on the community's flow of income' (ibid.: 19);
3. saved-up income is 'transferred' in time by banks so that it can be invested in the production of capital goods.

In the absence of banks, produced output would be immediately appropriated by producers and would consist of consumption goods only. As shown by Schmitt (1984a), if wages were paid out of material money owned by firms instead than out of nominal money issued by banks, output could never take the form of a stock for the economy considered

as a whole. This is so because in an economy deprived of banks saving can only be a zero-sum operation. Indeed, in such an economy saving is necessarily a credit transaction implying a lender and a borrower, where the lender is debited by the same amount the borrower is credited with. What the former loses is transferred to the latter, so that no net capital is formed for society as a whole.

[I]f an ordinary credit transaction takes hold of part of the income created by the emission of wages, it can consist, in the absence of banks' intermediation, in a contract agreed between an income holder who saves and a borrower, that is, of a symmetrical operation whose sum is nil for society.

(Schmitt 1984a: 163–4, our translation)

It is hard to know whether or not Keynes was perfectly aware of the implications of his intuition concerning the twofold nature, real and monetary, of capital. Be that as it may, it cannot be denied that the author of *A Treatise on Money* and *The General Theory* puts bank money and banks at centre stage and derives capital from the investment of saved-up income. 'We shall mean by the rate of investment the net increment during a period of time of the capital of the community' (Keynes 1930/1971: 114). It is macroeconomic capital (capital of the community) that Keynes wants to explain, and it is through saving and investment that current income can be transformed into capital.

The crucial message that we can derive from Keynes's argument is that the presence of instrumental goods is not enough to conclude to the existence of capital. Instrumental goods existed well before the creation of banks, yet physical means of production cannot straightforwardly be identified as capital goods. The financial component of capital is essential, and has been made possible by the discovery of double-entry bookkeeping and the creation of banks. What matters here is the possibility to establish a link, build a bridge, between present and future, and this can only be obtained through a financial intermediation allowing for the transformation of current income into a positive sum of net capital. '[A] society without capital could nevertheless be equipped with machines, even sophisticated instrumental means available to human labour. In fact, purely instrumental capital belongs to the prehistory of capital, which acquires its modern signification with the transformation of income: every capital *stricto sensu* is at the same time monetary and real' (Schmitt 1984a: 151, our translation).

Keynes's crucial contribution to the theory of capital is to have made it clear that this concept acquires all its significance with the introduction of banks. From then on, any analysis of capital has to account for both its monetary and real aspects and has to explain how a positive net capital can be formed out of saved-up income as well as how fixed capital can emerge as the result of a macroeconomic investment.

In order to verify if a causal link can be established between capital accumulation and economic crises, the positive analysis of capital must be accompanied by a critical analysis of the way capital actually accumulates. Among the authors considered in this section, Marx is the one who has developed the most the idea that capital accumulation will unavoidably lead, after a series of increasingly devastating crises, to the self-destruction of capitalism. Yet, his analysis rests on a labour theory of value that has proven wrong and cannot be retained. General equilibrium analysis does not fare better, and Walras's identification of capital with a productive service deprives this concept of its deep originality. Böhm-Bawerk's analysis is more interesting, but his roundabout methods of production do not provide any useful insight into the possible relationship between capital accumulation and macroeconomic crises. Keynes's contribution to this specific topic is also very limited, his concept of marginal efficiency of capital being in fact of a microeconomic nature. The lessons we can derive from the history of economic thought, from the Physiocrats to Keynes, are very important indeed, yet they pertain to the positive analysis of capital much more than to the analysis of its link to economic crises. In the next section we will complete this positive analysis and start considering whether capital accumulation can be at the origin of a structural pathology explaining the outburst of economic crises.

Capital formation as structural cause of economic disorder

Let us consider a modern monetary system. The first question we have to answer is how positive capital can be formed in such a system. The first elements to answer this question are provided by the economists of the past whose contribution we have just analysed: capital derives from saving and initially takes the form of a stock. Saving is a monetary concept related to income and the transformation of produced output into a stock calls for a positive investment. Income and real output appear therefore as the necessary prerequisite for any process of capital formation. Saving, lending, and investment are the operations carrying it out.

The initial step is production. As shown throughout this book, production is an instantaneous event, an emission through which physical output is given a monetary form and a positive income is deposited with banks. Are the saving of this newly created income and its lending on the financial market enough to transform it into capital? One could be tempted to answer yes, because such a financial transaction is enough to build a bridge between the present income of the saver and the future income of the borrower. Yet, this zero-sum transaction does not entail the formation of a positive stock, nor does it lead to a positive net loan. In order to be at the origin of a capital, the loan must be associated to the formation of income, that is, it must take place at the very instant income is formed as a bank deposit. In this regard, Schmitt (1984a) distinguishes between a quantum and an ordinary credit. The payment of wages is a quantum credit; its result is the formation of a new net income. The lending of an already-existing income, on the contrary, is an ordinary credit. Now, at a closer sight it appears that the emission of wages (a quantum credit) is always and necessarily accompanied by an ordinary credit. As the result of the payment of wages, income is formed as a bank deposit and is thus immediately lent. This is so because, according to double-entry bookkeeping, every amount entered on the liabilities side of a bank's ledger is also entered on its assets side. Hence, firms are immediately credited with the amount saved by workers, who maintain their right over produced output in the form of financial claims on banks. As soon as it is credited to firms, the income saved by wage earners is 'spent', that is to say, invested in the 'purchase' of current output. In fact, the transaction carried out by firms is not a final purchase, since workers are still the owners of the income lent to firms. The result of this first investment is therefore the formation of a stock that firms do not own although it is at their disposal. Saved from the very moment of its formation, income is immediately transformed into a capital, which is both financial (workers' claims on bank deposits) and real (the stock at firms' disposal).

In Schmitt's words,

[t]he emission [of wages] transformed into an ordinary credit gives rise, face to face, to products stocked (by firms) and (at the pole of income holders) the right to collect these stocks. Income is immediately destroyed and replaced by a net capital – financial (the claim on deposited incomes) and real (the stocks).

(*ibid.*: 161, our translation)

In this process, banks act as intermediaries and their presence is essential. Without them, firms would borrow the income saved by workers and no net capital could be formed. Indeed, in a personal relationship between lender and borrower, the former loses what the latter obtains and the result is a zero-sum transaction. When the payment of wages is carried out by banks, things change radically: what is saved by workers is not lost by them; income is replaced by capital and workers maintain their right over it in the form of financial claims on their bank deposits. What matters here is that workers' credit is balanced by the stock formed as a result of the investment carried out by firms. The object of workers' credit with their banks is the stock of produced output physically deposited with firms. The logical succession of events is therefore as follows.

- 1) Wages are paid to workers: banks debit firms and credit wage earners.
- 2) Income is formed and immediately deposited with banks.
- 3) Saved-up income is transformed into capital and lent to firms.
- 4) Capital is invested in current production and a stock is formed with firms.
- 5) Firms' debit with banks is balanced by the stock formed with them.
- 6) Workers' credit with banks defines their economic right over the physical output stocked with firms.

What we have just exposed is the quantum macroeconomic analysis of capital-time, the first capital formed in a monetary economy of production as a direct consequence of the passage of time. What has still to be explained is the process leading to the formation of fixed capital. How is it that capital-time is partially transformed into fixed capital? Being formed through saving, capital-time is perfectly reversible: when income holders (wage earners) give up their bank deposits to purchase current output, capital-time is destroyed and replaced by an equivalent income that, through its final expenditure, is also instantaneously destroyed. The formation of capital-time entails that of a positive stock; income holders' final expenditure defines the purchase of this stock. It is thus clear that if capital-time is to be partially transformed into fixed capital, part of the initial saving must become final, that is, income has to be definitively fixed into capital.

As in the case of capital-time, fixed capital is obtained through an investment. Yet, this time the relevant investment must imply the final expenditure of income and, as a necessary consequence, the final purchase of output. It is up to firms to carry out this investment. Now, firms'

investment can define a final purchase only if it implies the expenditure of their own income. In other words, firms' investment must be financed out of firms' profit (realized or anticipated). This necessary condition, however, is not sufficient to the formation of fixed capital. Indeed, if firms were to spend their profit on the commodity market, their final purchase of produced output would not substantially differ from the expenditure carried out by any other income holder on this market. Their investment would amount to the expenditure of a positive income in the purchase of consumption goods. If this were the case, production would consist of consumption goods only, the entire output being purchased through the final expenditure of wages carried out by consumers and firms on the commodity market. This means that, in order to entail the formation of fixed capital, firms' investment must define the final expenditure of firms' income (profit) on the *labour* market.

As a matter of fact, it is possible to maintain that fixed capital takes initially the form of capital-time, as profits are but that part of saved-up income that is transferred to firms on the market for produced goods and services. At the moment a profit is formed, firms become the owners of part of the initial stock defining the real content of the capital-time resulting from the payment of wages. 'The formation of profit determines the equivalent creation of a stock of wage goods. This stock is the first definition of fixed capital. [...] The first "content" of fixed capital is thus a collection of stocked wage goods, that is, a pure capital-time' (Schmitt 1984a: 170, our translation). The passage from this initial form to fixed capital proper is carried out by the investment of profit. As a matter of fact, only part of firms' profit is transformed into fixed capital. The part that is redistributed as interests and dividends is eventually spent in the purchase of stocked wage goods and is destroyed both as capital and as income. 'Things are clear: profit redistributed as non-wage income is destroyed in its double aspect: households consume (destruction of the financial capital) and, by so doing, take hold of an equal amount of stocks' (ibid.: 174, our translation). It is the part of profit that firms invest in the production of a new output that gives rise to fixed capital. The newly produced goods are purchased by firms at the very instant of their production: the sum invested is transformed into a financial fixed capital, while the output thus obtained makes up a stock of fixed capital goods that will never be purchased by households.

The initial stock of wage goods defining the first content of fixed capital is replaced by a stock of instrumental goods that firms obtain by investing their profit, that is to say, by spending it in the production of fixed capital goods. As shown by Schmitt's quantum macroeconomic

analysis, the investment of profit takes place on the labour market. In this case, workers' output is appropriated by firms since the payment of wages for the precise reason that wages are paid out of a positive income (firms' profit). In exchange, workers obtain a financial claim over the wage goods still owned by firms as a stock defining the real content of their profit. Firms' investment of profit on the labour market has the double result of creating a macroeconomic financial capital and a stock of fixed capital goods. It is thus confirmed that:

- a) saving is at the origin of capital – indeed, both capital-time and fixed capital derive from a saved-up income, the main difference being that capital-time is reversible and comes to an end when savings are spent, whereas the income transformed into fixed capital defines a macroeconomic saving, that is, a saving that will never be spent by consumers;
- b) time is crucial – it is because of time that a financial bridge is required to create a link between present and future: capital-time provides this link, while fixed capital derives from a transformation of capital-time (it is that part of capital-time that is fixed into instrumental goods);
- c) every capital is necessarily financial and real at the same time – '[d]uring a financial capital's life time, output is itself a capital; real capital is the logically-necessary "twin" of financial capital: each of the two capitals is the condition of existence of the other'.

(*ibid.*: 491, our translation)

So far we have devoted our attention to the positive analysis of capital formation. Our aim now is to show that the process of fixed capital formation as it actually occurs is a structural cause of economic disorder. To do so, let us reconsider what happens when firms invest their profit in the production of instrumental goods. As we have already noted, this investment defines the expenditure of a positive income (profit) within the payment of wages. It is precisely because wages are paid out of profits that firms appropriate fixed capital goods since their production. This is not what happens when wage goods are produced and that explains why workers own them, even though they are immediately transformed into a stock deposited with firms. The wage goods defining the real content of capital-time are still in the economic possession of income holders, who will eventually get hold of them at the moment they will spend their income on the commodity market. However, through the sale of their stock of wage goods at a markup, firms can also become income holders and, by the same token, the owners of part of the initial stock of wage

goods. When doing so, firms take the place initially held by workers. They take on the role of intermediary and are called upon to manage the income saved by their initial holders. Indeed, profit formation is the way through which part of national income is withdrawn from consumption in a monetary economy of production. To this extent firms are personalized: they act on behalf of households and people (shareholders) own them. The initial stock of wage goods simply changes hands in the same measure as the income initially formed through the payment of wages is transferred to firms via the commodity market.

If profits are redistributed as interests and dividends, and then spent for the purchase of consumption goods, the income saved in the form of profit is destroyed and the wage goods owned by firms are given up to shareholders and creditors. In this particular case, the totality of the capital-time initially formed to the benefit of wage earners and then partially transferred to firms is spent on the commodity market and destroyed, and households purchase the totality of the stock of wage goods defining its content.

Something rather different happens when, instead of being redistributed, profit is invested in the production of fixed capital goods. This time profit is spent within the payment of wages, which means that two transactions are simultaneously present within the same event. One transaction – the expenditure of profit – defines the purchase by firms of the newly produced capital goods; the other – the payment of wages – brings about the creation of new financial claims (claims on bank deposits) to the benefit of workers. Instead of becoming the owners of their own output – as it happens when wages are not paid out of profit – workers give up their financial right over fixed capital goods in exchange of equivalent rights over the wage goods still stocked with firms. Through the investment of profit, the capital-time owned by (personalized) firms and corresponding to a stock of wage goods is transformed into a fixed capital whose *real content* is a stock of fixed capital goods (Figure 7.1).

Figure 7.1 illustrates the process of fixed capital formation by distinguishing two phases and two periods. In period p_0 , only wage goods are produced and a profit is formed through their sale. The result of this first phase is the constitution of a stock of wage goods still unsold and owned by (personalized) firms. In period p_1 , firms' profit is invested in the production of instrumental goods. The payment of wages carried out in the second period includes therefore the expenditure of profit, and defines the purchase of fixed capital goods by firms. In Figure 7.1, we have represented this production only, leaving aside that of new wage

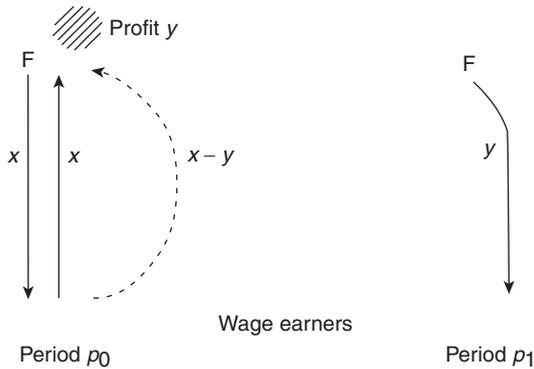


Figure 7.1 The process of fixed capital formation

or consumption goods in order to emphasize the crucial transaction, that is, the investment of profit.

What has to be clearly understood here is that nothing is wrong with the way profit is formed in period p_0 as well as with firms' decision to invest it in the production of instrumental goods. What has to be put under scrutiny is not this decision but its implications in the present system of payments or, more precisely, the implications of the way payments are carried out. Now, it is a matter of fact that today the investment of profit defines the expenditure of a positive income on the labour market. This means that, by paying wages out of profit, firms purchase labour itself. '[P]rofit is invested in a totally-original operation; it purchases labour; it is its remuneration. This time labour is itself a commodity; it is purchased by the transformation of profits into new wages. [...] In its investment, profit is a purchasing power over human labour' (Schmitt 1984a: 204, our translation). As the purchase of labour defines the purchase of its output, the investment of profit 'empties' workers' remuneration from its real content, which is definitively appropriated by firms. Every time wages are paid using nominal money, as in period p_0 , wage earners obtain what Schmitt calls 'full' money, that is, money (a numerical form) with a real content (currently produced output). By giving money a physical content, the payment of wages transforms nominal money into income and wage earners acquire the economic ownership over their output (they own it in the form of financial claims on bank deposits). By contrast, if wages are paid using a positive income, which is the case when profits are invested, wage earners' output is purchased by firms and the payment of wages brings them (workers) an 'empty' money. 'The emission of wages is *full* when the formation of

wages does not coincide with the expenditure of profits. On the contrary, the emission of wages is *empty* to the extent that it identifies itself with the expenditure of profits, a condition that is satisfied with the investment of profits' (ibid.: 205, our translation).

The presence of empty money is symptomatic of a pathology that introduces a structural disorder into the economic system. Indeed, the purchase carried out on the labour market leads to the economic expropriation of income holders (shareholders included) of fixed capital goods to the benefit of firms. What is worrying here is that the instrumental goods instantaneously withdrawn from wages are appropriated by firms that are depersonalized: 'The product withdrawn from money and households is *definitively appropriated by a "non-person", the disembodied set of a country's firms*' (ibid.: 208, our translation). The concept of disembodied or depersonalized firms may be difficult to grasp. Despite being very abstract (as is clearly conveyed by the choice of the words, disembodied firms are nowhere to be found as physical entities), this concept gives the correct idea of what happens when the investment of profit implies its expenditure for the purchase of labour. When output is purchased on the commodity market, it is always a person or a personalized institution that ends up owning it. Workers, pensioners, investors, shareholders, persons benefiting from social assistance, (personalized) firms and companies are all possible elements of the set of income holders and, as such, they can become the final owners of produced output. No structural disorder can arise from the expenditure of income on the commodity market. Pathology emerges only if part of current output is definitively lost to income holders, that is, if no one of them can get hold of it. This can be the case only when output is withdrawn from income holders at the very moment income is formed. When profit is invested, the output that should 'fill' the wages paid to workers is immediately 'fixed' into firms. As it never enters into workers' ownership, it will also never be obtained by any other category of income holders, personalized firms included. When their wages are paid out of profit, workers' income is in reality made up of empty money, and no matter how many times it changes hands and to whom it is transferred it will always remain empty money. The concept of disembodied expresses the fact that, because of the investment of profit, fixed capital goods never become the real content of money wages and, therefore, will also never become the real content of any other income.

At a superficial analysis it would seem that the appropriation of instrumental goods by disembodied firms has no serious consequence, because the loss suffered by workers producing them is compensated

by the stock of wage goods still available with firms. Indeed, the claims workers obtain in exchange of the fixed capital goods appropriated by disembodied firms give them the right to get hold of the wage goods produced in period p_0 and still unsold. Yet, this amounts to saying that households obtain nothing else than what they already owned in p_0 . On the whole, over the two periods households (and any other income holder that could take their place) are expropriated of the totality of fixed capital goods.

The concept of alienation was introduced into economics by the Classics, and Marx was the author who tried the most to show that alienation of labour is the unavoidable consequence of capitalism. His analysis rests on the labour theory of value, and it is by distinguishing between labour and labour-power that Marx introduces the elements enabling him to support his claim. In fact, Marx's theory is simple: labour-power is a commodity, and when workers are paid their wages, they sell it to firms, which can exploit it in order to derive a positive surplus-value. Alienation is therefore inherent in what Marx believed to be the very nature of capitalism: the purchase of workers' labour power. Since the publication of *Capital*, it has been proved that Marx's distinction is not only far-fetched, but is logically unacceptable, because it prevents the monetary formation of profit. It is with Schmitt's quantum macroeconomic analysis that alienation re-emerges as a consequence of the disorderly working of capitalism. This time the purchase of labour is not embedded in the nature of capitalism. Labour becomes a commodity only if profit is invested on the labour market, that is, only if the payment of wages is carried out through the expenditure of profit. In Chapter 10, we will point out that a capitalist system is possible, in which capital accumulation occurs without profit being spent on the labour market. In the systems of domestic payments capitalism has known so far this has never been the case. Yet, nothing but the ignorance of the logical laws of macroeconomics prevents the implementation of a monetary reform enabling the passage from a disorderly to an orderly system.

The existence of alienation to the profit of disembodied firms might seem a philosophical quibble with no empirical consequences. After all, workers producing fixed capital goods can spend their empty money in the purchase of the wage goods stocked as profit in period p_0 . On the whole, households purchase the totality of wage goods and firms get hold of instrumental goods. Does it really matter whether capital goods are 'fixed' with disembodied firms or owned by personalized firms? The answer is yes and in the next section we will explain why.

Inflation and unemployment as twin outcomes of the same cause

Fixed capital goods formed in period p_1 will be used in the following periods as means of production, thus raising considerably the physical productivity of labour. Yet, their use entails their wear and tear and calls for their amortization. Now, from a macroeconomic point of view, amortization implies the production of new instrumental goods replacing those used-up in the production of final output. For purely didactical reasons, let us suppose that fixed capital goods produced in period p_1 are entirely worn out at the end of period p_2 . In this case, amortization has to take place also in period p_2 : the output produced in p_2 is therefore made up of new wage and investment goods and of amortization goods (Figure 7.2).

In Figure 7.2, we follow Schmitt's method, distinguishing the set of firms producing amortization goods from those producing wage and investment goods. At this point it is crucial to understand that the correct analysis of what happens in p_2 depends on what happened in p_1 . The nature of the fixed capital formed in p_1 is decisive to grasp why the income formed in the production of amortization goods cannot be spent in their final purchase. If personalized firms owned fixed capital, amortization goods would initially be owned by workers and eventually by shareholders. The problem is that in p_1 income holders are expropriated of fixed capital to the benefit of disembodied firms. As a consequence, the production of amortization goods is, from the outset, a production devoted to the maintenance of the fixed capital pathologically formed in p_1 . Part of the working population works for disembodied firms in order to compensate for the loss suffered by pathological capital because of wear and tear.

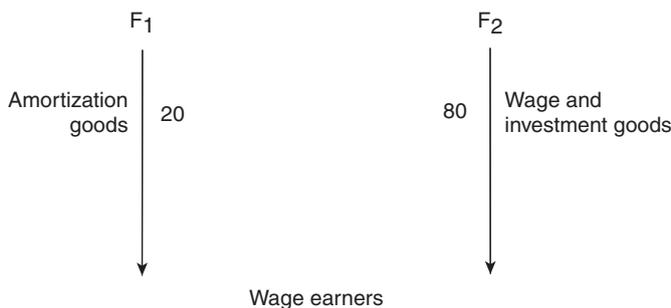


Figure 7.2 The production of amortization goods

It is true that the population benefits from the increased productivity made possible by fixed capital, so that it might be reasonable to consider the entire process as a fair exchange between the work done in favour of fixed capital and the increase in use-values obtained through the use of fixed capital goods. Once again, however, things are different from what they appear at a first glance. More precisely, the exchange we have just described is only part of the whole story, which unfortunately develops along a less favourable path.

As amortization goods are doomed to maintain fixed capital's original value and as in the present disorderly system disembodied firms own fixed capital, the income earned in p_2 by workers producing amortization goods can only be spent for the purchase of what is produced by the second set of firms. What is true for workers employed by firms F_1 is also true for those employed by firms F_2 and for any other economic agent taking their place as income holder: being produced to the benefit of pathological capital (disembodied firms), amortization goods can never be purchased by income holders. As illustrated by Figure 7.3, this implies that the totality of income formed in the system is necessarily spent for the purchase of F_2 's output.

Firms F_2 sell their products, whose production costs are equal to 80 money units, for 100 money units, and realize therefore a profit equal to 20 money units (the difference between F_2 's inflows and outflows). What has to be emphasized here is that these 20 units of profit realized by F_2 are a kind of extra profit of pathological origin. Whatever be the decisions taken by income holders, their demand (100) is greater than the supply they have access to (80). From this it follows that, independently of income holders' behaviour, F_2 's profits are increased by 20 money units.

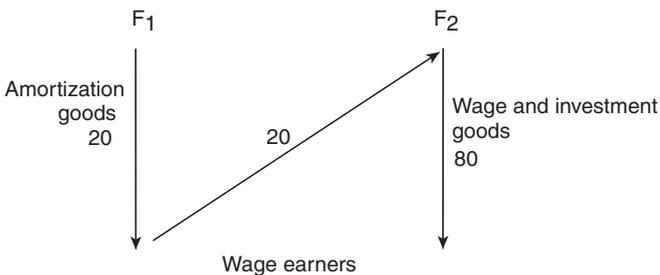


Figure 7.3 The expenditure of the income formed in the production of amortization goods

The next step consists in investigating the logical consequences of this unavoidable extra profit over the production of F_2 . At the beginning of our analysis we did not specify the nature of F_2 's output beyond the fact that it is made up of investment and wage goods. Analysis provides us now with a piece of information enabling us to infer that, from the outset, 20 out of the 80 money units paid as wages by firms F_2 have their origin in F_2 's profit. In other words, F_2 's production includes an investment of profit equal to 20 money units. The appropriate analysis here is exactly the same that we applied to period p_1 : the payment of wages out of a positive profit leads to the formation of a pathological fixed capital. Of the 80 money units of output produced by F_2 , 20 consist in profit goods disembodied firms appropriate in the payment of wages. Given that the investment of profit defines its expenditure on the labour market, investment goods are taken out from wages at once or, more precisely, they never became wages' real content. The purchase of labour itself empties wages of their content, so that workers or, more generally, income holders obtain *de facto* an empty money (Figure 7.4).

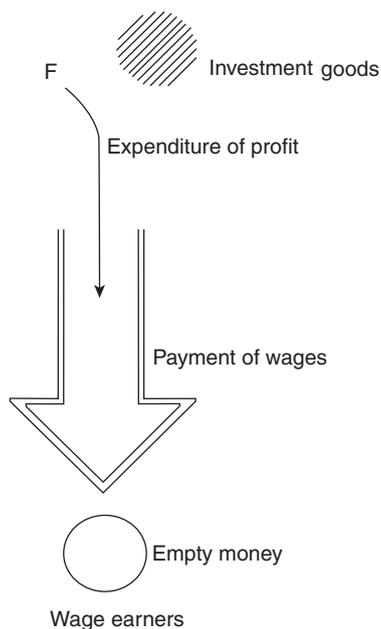


Figure 7.4 The investment of profit

As claimed by Schmitt (1984a: 219, our translation), ‘part *a* [20] of wages *a*’ [the total wages – 80 – paid by F_2] is not an income: workers obtain it directly *in the form of a financial claim*. The object of this claim is well known: amortization goods’. Albeit households do not purchase amortization goods, they eventually pay for them. Yet, this does not modify the fact that, emptied of their real content, wages paid for the production of profit goods cannot be ‘filled’ with any other product. Unlike what happens in period p_2 when fixed capital goods are initially produced, no wage goods are left over from previous periods, so that the formation of empty money cannot be ‘neutralized’ in its pathological effects. In the words of Schmitt, not only wages paid for the production of new profit goods ‘are empty with respect to newly produced output, but they are also empty with respect to any previous output (wage goods); they merely allow income holders to pay for amortization goods’ (ibid.: 223, our translation) (Figure 7.5).

Figure 7.5, derived from the one drawn by Schmitt (1984a: 220), illustrates the entire process occurring in period p_2 . As can easily be observed, production of amortization goods unavoidably gives rise to an equivalent production of new profit goods. This is what Schmitt calls dual production, a phenomenon explaining that, in its pathological form, the process of capital accumulation degenerates into a process of capital over-accumulation. In our numerical example, instead of maintaining fixed capital at the initial level of 20 money units, amortization elicits a new dual production increasing it further by 20 money units. The value lost by fixed capital because of wear and tear (20 money units)

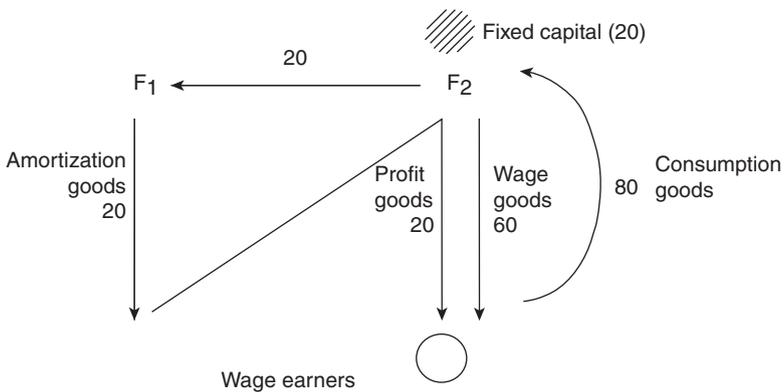


Figure 7.5 The production of amortization goods and its consequences

is compensated by amortization and, at the same time, a new investment of profit (also equal to 20 money units) brings its total amount up to the level of 40 money units. Schmitt's conclusion is dismaying: '[t]o the extent of dual production, workers do not produce for income holders; they do not produce for people: *they are enslaved to Capital*' (ibid.: 223, our translation). The capital Schmitt refers to is the pathological capital formed through the investment of profit and whose existence is at the origin of both inflation and involuntary unemployment. Alienation is not the result of labour-power's exploitation, nor is it due to entrepreneurs taking advantage of workers in a class struggle opposing people to people. The pathology described here is structural: it is caused by a lack of consistency between the present system of domestic payments and the logical laws of monetary economies. What is required to get rid of pathological capital is therefore not a political but a conceptual revolution. What is urgently needed is a greater intellectual effort, not a bloody civil war.

The first consequence of the empty money formed in period p_2 is inflation. Indeed, it is immediately clear that the existence of 20 units of money with no real content whatsoever inflates the number of money units with respect to the output still available in the market. Already purchased when profit is invested, profit goods are no longer part of global supply. On the other hand, albeit deprived of their real content, the wages paid to workers producing profit goods increase global demand. This increase is purely nominal, but this is precisely what characterizes inflation: a nominal increase in global demand entailing a loss of purchasing power for every money unit. In our example, total output – equal to 100 money units – is reduced to 80 units by the purchase of profit goods on the labour market, while the total amount of money units is still equal to 100. The distribution of a product initially carried by 80 units over 100 money units reduces the real content of each money unit.

Unlike what happens when the nominal excess demand is due to an excess of credit, the increase in nominal demand caused by the empty emission defining the investment of profit is never compensated. Since amortization of fixed capital engenders an equivalent dual production of profit goods, the economy is split into three sectors: the first sector produces wage goods, the second amortization goods, and the third profit goods. Profits being derived from wages and the production of amortization goods diminishing that of wage goods, the pathological production of profit goods in the form of new capital goods (over-accumulation) can never increase beyond that of wage goods. In other words, the limit of the pathological process of capital accumulation and

over-accumulation is reached when production is equally distributed in the three sectors, each sector employing one-third of the working population (not expressed in number of persons but in wages, because, let us recall it, the wage unit is the specific unit of measure in economics). As explained by Schmitt, this limit is set by the maximum level amortization can reach, given that 'the production of amortization goods is duplicated twice in the period [p_2]' (ibid.: 225, our translation). Indeed, the production of amortization goods – 20 money units in our example – determines the amount of wage goods purchased by workers employed to produce amortization goods and, at the same time, the amount of profit goods resulting from the investment of profit.

Before the limit is reached, capital accumulates at an increasing pace and so does the empty emission defined by the production of extra profit goods (in the form of new fixed capital goods). Inflation is the by-product of over-accumulation and this explains why, despite the enormous increase in physical productivity, consumer price indices have continued to rise in the past decades. Yet, as economies approach the limit of capital accumulation another problem sets in, leading to deflation and involuntary unemployment. The problem is that capital has to be remunerated and amortized.

Let us go back to our numerical example and suppose that we have reached, in period p_n , a level where production of amortization and extra profit goods is equal to 30 wage units whereas the production of wage goods is of 40 wage units. Let us assume also that the capital accumulated until period p_{n-1} is equal to X and that, if it were to increase from X to $X + 30$ at the end of p_n , the production of amortization goods in the following period, p_{n+1} , would not be enough to reproduce the loss it suffers because of wear and tear. Alternatively put, we suppose that the limit of capital expansion is reached in p_{n+1} , so that from this period onwards the production of amortization goods cannot be greater than $33\frac{1}{3}$, and that the amortization of a capital equal to $X + 30$ requires indeed more than a production of amortization goods equal to $33\frac{1}{3}$. This clearly means that the production of profit goods in the form of new fixed capital goods has to be reduced since period p_n in order to limit the increase of total capital. In his 1984, book Schmitt observes that such a result can be obtained by investing part of the extra profit either on the financial market or on the production of new profit goods in the form of extra wage goods. 'When profit-dividends can no longer rise in front of the persistent increase of capital, firms can only choose between two alternatives: to reduce the production of profit-goods or to maintain it, but asking workers to produce wage-goods' (ibid.: 237, our translation).

While the lending of part of the extra profit leads immediately to unemployment and to an increase of households' indebtedness, the production of extra wage goods leads first to deflation. Why is it so? The answer is straightforward: the increase in the supply of wage goods is not matched by an increase in available income. The wages paid to workers producing extra wage goods having been 'nourished' by profit, no positive income is formed by what we have called, following Schmitt, dual production. Thus, whereas the newly-produced wage goods inflate supply, demand remains at its previous level: a gap appears between total supply and total demand, a gap that defines a situation of deflation. 'The vice is patent: [extra] wage-goods must be sold a second time, that is, on the commodity market, after having already been sold on the labour market. The domestic economy provides the necessary purchasing power *only once and not twice*' (ibid.: 238, our translation). The consequences of this pathological increase of total supply are easy to forecast: firms, faced with the impossibility to sell the totality of their output, will first reduce its selling price and then lay off part of their workers.

Quantum macroeconomic analysis shows that the pathological formation of fixed capital occurring in p_1 is at the origin of both an inflationary and a deflationary gap between total demand and total supply. This might sound as an astonishing and contradictory claim, because it seems inconsistent to maintain that global demand can be greater and lesser than global supply at the same time. Economists have attempted in vain to reconcile the apparent opposition between theory – which they believed to exclude the coexistence of the two disorders at the national level – and facts – which have forced them to recognize the existence of this possibility by describing it as stagflation. Indeed, it is only by recurring to quantum analysis and rejecting the principle of *tertium non datur* that the contradiction can be avoided. In period p_2 the investment of profit is an empty emission that creates a numerical difference between total demand (120 money units) and total supply (100 money units). Indeed, the demand exerted by firms on the labour market (20 money units) adds up to that exerted by workers out of their wages (100 money units). The increase in total demand is purely nominal, since workers employed in producing profit goods are paid in empty money. When, in period p_n , part of profit goods consists in extra wage goods, the gap between global demand and global supply arises again. Once again, global demand increases because of the investment of profit, while global supply is still equal to 100 money units. However, this time a numerical gap appears also between global supply ($100 + y$)

and global demand (100), where γ is the part of profit goods having the form of wage goods and global demand is measured in terms of income. Even though they have already been purchased by firms since their production, goods γ are supplied again on the commodity market, where they increase supply nominally. Here too the presence of a nominal difference is symptomatic of a pathological disorder, which is the very nature of deflation.

Inflation is a nominal difference created when global demand is numerically increased (that is, inflated) with respect to global supply, while deflation is a nominal difference appearing when global supply rises numerically with respect to global demand. Inflation is formed when the rise in demand is exerted out of empty money; deflation arises when supply is partially 'duplicated'. In the case of inflation, the excess demand is a 'false' demand, which is not backed by a positive income. In the case of deflation, excess supply is a 'pathological' supply, since it corresponds to the presence on the commodity market of goods that have already been subject to a final purchase on the labour market. Despite being pathological disorders of opposite sign, these two anomalies do not cancel one another. The excess demand defining inflation cannot counterbalance the lack of demand defining deflation, since what is lacking in the case of deflation is an amount of income and not of empty money. Analogously, inflation can be avoided only by reducing total demand to the level of actual supply, and not by increasing supply in a purely nominal way.

Even though the decision to invest extra profit in the production of wage goods does not reduce the employment level, the ensuing deflation will eventually force firms to stop part of their activity, thus increasing involuntary unemployment. Once the limit of one-third of national income has been reached by the three sectors, unemployment is bound to grow to the extent that the sector producing profit goods shrinks. The maximum level of involuntary unemployment is reached when one-third of the working population is unemployed. There is no need to stress how dramatic it would be if this limit were attained and how urgent it is to find a way to avoid this apparently unavoidable fate. In Part III, we will see that a reform exists that, if implemented, would enable to replace the present, disorderly system with an orderly one, in which both inflation and involuntary unemployment would be absent. For the time being, let us consider the role played by interest rates in the spread and intensification of economic and financial crises.

8

Interest, Rate of Interest, and Crises

Interest and capital are two closely related topics that have been widely discussed by economists. Although in today's economic literature interest is mostly identified with contract-interest on loans and its rate with the market rate of interest, the problem of its origin has still to be satisfactory solved. Economists of the past struggled to argue the existence of an original interest enabling that of the interest on loans, where the latter is essentially nothing but the effect of the former. They all agreed that capital is at the origin of interest, yet they disagreed about the kind of causal link that exists between these two concepts. Is capital a direct or an indirect cause of interest? The question whether or not capital is a macroeconomic factor of production is still open today and, despite appearances to the contrary, still a subject of controversy. Indeed it is only thanks to Schmitt's quantum macroeconomic analysis that a final answer is possible, the payment of human labour being the only transaction capable of transforming money into income.

In this chapter, we first review some of the main contributions of the past to the analysis of interest. The role played by capital and time is an essential part of these contributions and so is the debate about the relevance of physical productivity. What some of the greatest economists of the past teach us is that interest derives from capital and time, even though capital cannot be considered as a factor of production on a par with labour. They also distinguish this (original) interest from the interest on loans, which is considered a subsequent and logically secondary consequence of the former.

Keynes's contribution differs from that of his predecessors mainly in that he does not consider the rate of interest as the balancing factor allowing for the equality between macroeconomic saving, S , and investment, I . He claims, on the contrary, that the relationship between S and

I is an identity and thus sets a new logical framework within which interest has to be explained. For that reason, a section of this chapter is devoted to Keynes's contribution and to a critical appraisal of that of his followers.

Capital is formed through the investment of profit and defines a macroeconomic saving: this is what Schmitt (1984a) derives from Keynes's identities, both between global supply, *Y*, and global demand, *C + I*, and between *S* and *I*. A correct analysis of interest has to respect these identities and show how it is possible to derive a positive macroeconomic value from capital given that labour is the sole macroeconomic factor of production. This is one of the main arguments developed in this chapter, an argument that leads to the analysis of the impact of interest rates on economic and financial crises. The last section of this chapter deals with this analysis and shows that capital accumulation tends to reduce the rate of profit of the economy taken as a whole, a tendency that, by narrowing the gap between the rate of profit and the market rate of interest, creates the conditions for an economic crisis bound to hamper economic growth to an increasing extent. As more and more capital is accumulated, it becomes increasingly difficult to remunerate it and the process of capital accumulation has to slow down. As the macroeconomic rate of profit gets closer to the market rate of interest, productive investment must be reduced, which has a negative effect on employment. The economic crisis engendered by the impact of capital accumulation on profit and interest rates is then worsened by the financial crises induced by the presence of a growing pathological capital and of the speculative transactions it feeds.

The theory of interest: Contributions of the past

Capitalization versus physical productivity

Let us attempt to clarify some of the main contributions to the analysis of interest provided by authors of the past by referring essentially to Böhm-Bawerk's and Fetter's contributions to the so-called capitalization approach. The views of these two authors are indeed particularly apt to emphasize the role played by time and capital in the determination of interest as opposed to that attributed to physical productivity.

Let us start with time. Both Böhm-Bawerk and Fetter recognize the centrality of time. The Austrian economist was arguably one of the first to develop a theory of interest in which time is the crucial element in the process leading to an increase in the macroeconomic income generated by production. According to his analysis, a difference in value

between present and future goods of the same kind can be explained by two reasons: on the one hand, psychological time preferences induce agents to underestimate their future needs, which they feel to be able to satisfy more easily than their present needs; on the other hand, higher physical productivity of what Böhm-Bawerk calls the roundabout methods of production will provide more output in the future with respect to the present. As for Fetter, he also puts time preference or time value at the core of his analysis, but refuses to consider physical productivity as a possible, even a secondary, cause of interest and rejects Böhm-Bawerk's explanation based on the roundabout methods of production.

As clearly put forward by Hülsmann (2002), Böhm-Bawerk's theory of time preference does not stand up to a critical examination, first because it is 'difficult to reconcile with the fact that values and prices are manifest in human choice and that choice is free' (p. 82), and secondly because present and future goods cannot be taken to be homogeneous so that no premium on the value of a present over that of a future good of the same kind can be postulated: 'one cannot even make claims on behalf of present and future goods "of the same economic quality" without contradicting oneself' (p. 82). However, Böhm-Bawerk's main contribution to the theory of interest does not consist in his analysis of psychological preference but rather in his intuition that interest:

1. defines an increase in value related to capital and time;
2. does not derive directly from physical productivity;
3. has to be distinguished between an original interest, *Urzins*, and an interest on loans, *Leihzins*.

Despite the fact that his analysis is in many respects neoclassical and marginalist, Böhm-Bawerk does not share the widespread belief that capital is a factor of production on a par with labour and land, and that interest is simply the income generated by this specific factor: 'he also fought the idea that "physical" capital is a distinct factor of production, capable of being treated on the same plane with the "original" factors, labor and natural agents' (Schumpeter 1954/1994: 901). His idea is that fixed capital allows for an increase in the value of produced goods owing to the longer period of production involved by its use. Böhm-Bawerk's concept of roundabout methods of production describes a situation in which final output is produced by labour with the aid of fixed capital goods. Now, instrumental goods have themselves to be produced and this requires the previous formation of a subsistence fund for supporting workers charged to fulfil this task. Capital itself is therefore derived

from labour so that, finally, the increase in value attributed by Böhm-Bawerk to the existence of the roundabout methods of production is brought about by a combination of human labour, past and present, and time.

Böhm-Bawerk's theory of capital and interest has been interpreted and criticized in many different ways, and our reading of his contribution makes no pretence to be more faithful to his real intuitions than the others. Our aim is merely to show that, interpreted according to the principles of quantum macroeconomics, his analysis provides the logical foundations for an up-to-date satisfactory explanation of the nature of interest. In this respect, one of the relevant features of his approach is the refusal to consider capital as a factor of production while recognizing that its presence and its 'time dimension' have an impact on the value of output. Another central aspect is the claim that this increase in value defines the true, original interest (*Urzins*) and that the interest on loans (*Leihzins*) is merely a consequence of the former, whose presence is the necessary condition for the existence of the latter.

Fetter is another economist whose analysis of interest is worth mentioning alongside that of Böhm-Bawerk. Indeed, he shares Böhm-Bawerk's idea that the *Urzins*, which he calls time-value interest, is logically prior to the interest on contract loans, the *Leihzins*, and links the former to the process of capitalization. According to Fetter (1914), capital is the fundamental cause for the existence of interest and the difference in the valuation between present and future goods is the cause of capitalization, that is of 'the process of putting a present worth upon any durable source of wealth and thus discounting its future uses by the act of exchanging it for other things' (p. 76). The similarity between the two authors' analysis is evident: both of them derive the existence of interest from that of capital and they both assume that the centrality of time is related to the existence of a rate of premium and discount between present and future goods. The psychological evaluation of goods is an important common feature of the analysis developed by the two authors who, in this respect, clearly belong to what was to become known as the neoclassical school. Their opinions diverge only slightly with regard to the explanation of the way a discount is formed between the value of present and future goods. According to Böhm-Bawerk, roundabout methods of production play as central a role as the individuals' tendency to underestimate the value of future goods, whereas Fetter sees time preference as the main cause of interest in a capitalization theory where interest is explained in purely psychological terms. As claimed by Hülsmann (2002), 'despite various disagreement with Böhm-Bawerk, all

later champions of time preference theory [...] would also agree that interest was a value differential between a “sooner” and a “later” good’ (p. 78). Yet, what we want to emphasize here is not the psychological component of the time preference analysis of interest advocated by Böhm-Bawerk and his colleagues, but their rejection, clearly stated by Böhm-Bawerk himself and by Fetter (1914), that capital’s productivity is the direct cause of interest.

Let us consider more closely Böhm-Bawerk’s distinction between *Urzins*, which we can call macroeconomic interest, and *Leihzins* or microeconomic interest, that is, the interest paid on loans. What Böhm-Bawerk and Fetter want to make clear is that the *Leihzins* would not exist if it were not founded on the *Urzins*, that is, that the formation of a positive macroeconomic interest is logically prior to that of its distribution among economic agents. Both authors do not deny the existence of interest on loans, but they reject the idea that interest can be formed simply through the imposition of a positive rate of interest on the money market. Like any other fundamental concept, interest has to be explained starting from its very *raison d’être* and cannot merely be postulated on the ground that its existence is empirically evident. A truly scientific approach requires an investigation endeavouring to determine the origin of interest, and both Böhm-Bawerk and Fetter are well aware that it cannot be found on the side of the *Leihzins*. Thus, for example, Fetter (1914) distinguishes the contract interest (the microeconomic interest) from what he calls the time-value (the macroeconomic interest) and writes:

Seeing the two problems [the determination of contract interest and that of time-value] as in large measure distinguishable, and seeking for the logical starting point in the study, I asked: Which of these two questions was prior in history and which is primary in logic? In both cases the answer was time-value. (ibid.: 76)

But our two authors go further in their analysis and reject also the widespread belief that interest is the direct result of the physical productivity of capital. In a period when economists were enthusiastically endorsing Walras’s neoclassical general equilibrium approach and the assumption, essentially microeconomic, of a plurality of factors of production (traditionally labour, land, and capital), to maintain that interest is not directly generated by capital is the mark of a great intellectual independence and acuity. We may not agree with the arguments put forward by Böhm-Bawerk and Fetter to prove their intuition, but we must recognize that they were able to clearly see the close relationship

existing between interest and time, and the logical impossibility to derive interest from the physical productivity of capital. As the following question asked by Fetter (1914) clearly suggests, the adjustment of prices is not the cause but the effect of interest, whose existence cannot be explained by considering capital as a factor of production. 'May we not then conclude that the cost-of-production-of-capital explanation of interest is a partial glimpse of an intermediate and subordinate process of the adjustment of prices, in part a mistaking of effect for cause?' (ibid.: 88).

Fisher is another author whose analysis of interest is often referred to by economists investigating Böhm-Bawerk's contribution. His point of view differs from Böhm-Bawerk's and Fetter's in that he rejects the idea that interest derives from the relative preference for present over future goods, claiming that the comparison between present and future values is itself dependent on the rate of interest.

When we say that interest is the premium on the value of a present house over that of a future house we are apt to forget that the value of each is itself based on a rate of interest. [...] In the process of discounting there lurks a rate of interest.

(Fisher 1907: 91)

If Fisher's claim that the money rate of interest enters into the determination of the value of goods, present and future, were correct, the view that interest is due to a difference in the time evaluation of similar goods would have to be rejected as (viciously) circular. Yet, the origin of interest would remain mysterious, the simple observation that an interest on contract loans has to be paid by borrowers once loans fall due being still in need of an explanation. In Fisher's analysis, Böhm-Bawerk's distinction between *Urzins* and *Leihzins* is blurred. In some of his writings, Fisher (1907, 1911) rejects Böhm-Bawerk's claim that interest is generated by the superior productiveness of roundabout processes of production and considers the proposition that interest is due to the productivity of capital as 'attractive, but [...] superficial' (Fisher 1911: 383). But then he modifies his judgement and, accepting Böhm-Bawerk's point of view, distinguishes the Austrian economist's contribution from the ordinary productivity theories: 'every one who has read Böhm-Bawerk should believe that *the ordinary*, or as Böhm-Bawerk calls them, the "*naïve*" productivity theories are snares and delusions' (Fisher 1913: 617). All in all, Fisher hesitates between rejecting the idea that interest is generated by physical productivity – 'whatever element is responsible for the existence of interest in the actual world, that element cannot be

physical-productivity' (Fisher 1907: 22) – and accepting time preference as 'the central fact in the theory of interest' (ibid.: 88) and making time preference itself depend on the rate of interest – 'If, therefore, we undertake to make the rate of interest depend on the relative preference for present over future houses, we are making it to depend on two elements in each of which it already enters' (ibid.: 91).

It seems possible to argue, therefore, that the crucial point that distinguishes Böhm-Bawerk's and Fetter's analysis from Fisher's is that the former two authors look for a logical explanation of the very existence of interest, while the latter seems to be merely interested in the mathematical determination of a rate of interest that is supposed to be somehow related to the productivity of capital.

To conclude this first introductory investigation of the contribution of past authors to the analysis of interest, let us briefly refer to Seager's (1912) critical appraisal of Fisher's (1911) article on the impatience theory of interest. In his contribution published by the *American Economic Review*, Seager (1912: 842) claims that there is a 'necessary or logical connection between physical-productivity as a general phenomenon of capitalistic production and value-productivity'. The problem is central and has to be rigorously examined while being aware that it cannot be solved simply by assuming that capital is a factor of production. Indeed, while it is undisputable that capital increases physical productivity, it is not at all evident that it is also at the origin of value. As intuited by Böhm-Bawerk and definitively established by Schmitt's quantum macroeconomic analysis, the truth is that only human labour is 'productive' in terms of value, so that if a logical relationship exists between capital physical productivity and value, it has to be explained by respecting the logically prior role played by labour.

Seager himself seems close to a correct understanding of the solution when he states that 'an increase in the total value-product as a consequence of the assistance which capital renders to production seems to me to follow as a logical necessary consequence' (ibid.: 843). To be true, an isolated quotation is clearly insufficient to suggest that Seager was aware of the fact that, although it is not directly at the origin of a positive economic value, capital allows for an increase in the value produced by labour. The idea that an increase of the physical productivity of labour made possible by the assistance of fixed capital is transformed into an increase of the value generated by labour was alien to Seager, and it is only with Schmitt's development of quantum macroeconomics that interest is finally understood in its very essence: a macroeconomic income derived from labour compensation.

Before explaining how it is possible to reconcile the existence of a unique factor of production, to wit, labour, with the macroeconomic nature of interest, let us spend a few words about Keynes's contribution to the analysis of interest.

Keynes's theory of interest

Keynes (1936/1946) rejects the conventional idea that interest payments correspond to the production due to fixed capital. He notably presents his own general theory of the rate of interest in Chapter 13 of his *General Theory*. In fact, there are three different theories of the rate of interest in that chapter. They cannot be reconciled, however, because they are contradictory.

According to the first theory proposed by Keynes in this regard, interest is the price of savings: 'the schedule of the marginal efficiency of capital must be said to govern the terms on which loanable funds are *demand*ed for the purpose of new investment; whilst the rate of interest governs the terms on which funds are being currently *supply*ed' (Keynes 1936/1946: 165, our emphasis). There would thus be, on one side, a demand for fixed capital (new investment) and, on the other side, a supply of financial capital (loanable funds). The demand for fixed capital would be a function of its marginal efficiency, whilst the supply of savings would depend on the interest rate. Hence, interest would not originate in the production process (that is, as an income formed owing to fixed capital invested in the real sector), but in the purely financial sector. Keynes justifies this theory referring to agents' liquidity preference.

Now, if firms demand new capital goods (according to their marginal efficiency), one has to consider also that these goods are supplied by income holders logically. These savers indeed supply 'loanable funds' (a financial capital), which firms demand on the relevant market. By deciding to consider only the demand for capital goods and the supply of financial capital, Keynes separates the marginal efficiency of capital from the rate of interest. However, if one considers the 'supply and demand of fixed capital and financial capital' altogether, one can understand that, in reality, the marginal efficiency of capital and the rate of interest are one and the same magnitude, the former being nothing else than another expression for the latter. If so, then the alleged dichotomy between fixed capital and financial capital – abstracting from any pathological working of our economic systems – is a figment of the imagination. In an orderly working economic system, in fact, fixed capital and financial capital are twin aspects of the same reality, since

the set of savers owns the set of fixed capital used by firms during the production process.

Now, after having maintained that interest is the price paid to savers, Keynes rejects this theory, claiming that '[i]t should be obvious that the rate of interest cannot be a return to saving or waiting as such. For if a man hoards his savings in cash, he earns no interest, though he saves just as much as before' (Keynes 1936/1946: 166–7). In this connection, Keynes's argument relies on empirical observation: like sight deposits with banks in a number of countries, some kinds of savings are not remunerated. In fact, according to this second theory of interest proposed by Keynes, (the rate of) interest is the price for disposing of liquidity (or the reward for abstaining from hoarding): 'the mere definition of the rate of interest tells us in so many words that the rate of interest is the reward for parting with liquidity for a specified period' (*ibid.*: 167). The so-called liquidity preference would induce thereby a demand for money that agents want to keep as a liquid store of value.

Indeed, Keynes refers to liquidity preference to show that (the rate of) interest is not the supply price of savings actually: '[t]hus the rate of interest at any time, being the reward for parting with liquidity, is a measure of the unwillingness of those who possess money to part with their liquid control over it' (*ibid.*: 167). Holders of 'liquid savings' (to wit, savings in money form) would receive no interest, because according to Keynes the latter is 'the reward for parting with liquidity'. In fact, no savings can be liquid logically: any savings are in fact a capital (hence something 'illiquid'), so that the payment of interest is due to each form of savings. In fact, when he refers to money's weak elasticities of production and of substitution, Keynes (*ibid.*: 241, footnote 1) notes that

[t]he attribute of 'liquidity' is by no means independent of the presence of these two characteristics. For it is unlikely that an asset, of which the supply can be easily increased or the desire for which can be easily diverted by a change in relative price, will possess the attribute of 'liquidity' in the minds of owners of wealth. Money itself rapidly loses the attribute of 'liquidity' if its future supply is expected to undergo sharp changes.

The third Keynes explanation of the rate of interest refers to the price equilibrating money demand and supply: '[t]he rate of interest is not the "price" which brings into equilibrium the demand for resources to invest with the readiness to abstain from present consumption. It is the

“price” which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash’ (ibid.: 167). In this view, money would therefore be a stock of purchasing power, similar to the so-called money balances in neoclassical economics. The rate of interest would therefore influence the quantity of money that economic agents want to keep in their own portfolio. According to Keynes (ibid.: 167),

if the rate of interest were lower, *i.e.* if the reward for parting with cash were diminished, the aggregate amount of cash which the public would wish to hold would exceed the available supply, and [...] if the rate of interest were raised, there would be a surplus of cash which no one would be willing to hold.

Now, it is surprising that in a volume aiming at explaining the working of a monetary production economy the author introduces money only after 170 pages as Keynes did: ‘[w]e have now introduced money into our causal nexus for the first time, and we are able to catch a first glimpse of the way in which changes in the quantity of money work their way into the economic system’ (ibid.: 173). In fact, there is a worse critique in this regard, referring to the way in which Keynes introduces money in that framework:

[l]iquidity-preference is a potentiality or functional tendency, which fixes the quantity of money which the public will hold when the rate of interest is given; so that if r is the rate of interest, M the quantity of money and L the function of liquidity-preference, we have $M = L(r)$. This is where, and how, the quantity of money enters into the economic scheme.

(ibid.: 168)

In this framework, the causality would run from the rate of interest to the quantity of money: economic agents would thus decide about the quantity of money to hold as a ‘liquid store of value’ depending on the rate of interest (exogenously given). This is tantamount to assuming that those units of money that agents dispose on any markets – as a result of an increase in the rate of interest – are lost for the economy as a whole. In fact, it is plain that, according to the net-asset conception of money adopted by Keynes in this framework, an agent’s expenditure defines another agent’s receipt. As a result, what an agent spends is earned by some other agent, and it is therefore impossible that such a process may change the quantity of money in the whole economic system. All in all

we need an essentially different theory of money and interest to explain both of them at macroeconomic level. Let us investigate this issue next.

The quantum macroeconomic analysis of interest

Let us start from the distinction between macroeconomic and microeconomic interest. The logical order of priority between these two forms of interest goes from the former to the latter, which means that before analysing how interest is distributed among economic agents it has to be explained how a positive macroeconomic interest can exist in the first place. The difficulty of this explanation lies in the fact that no macroeconomic income can be directly formed by capital, which is by no means a macroeconomic factor of production. To maintain that interest results from the use of capital much in the same way as wages result from the employment of labour makes it extremely easy and straightforward to explain interest. However, this apparent shortcut to scientific truth is totally inadequate, because it is only through the payment of wages that physical goods can become the object of a financial deposit with banks and acquire a monetary form. Wages are the only income originally formed in any economic system and all the other incomes must be derived from them. The payment of interest implies the pre-existence of a positive income, whereas the payment of wages is, alone, at the origin of income. If the payment of wages were to define the final purchase of the goods and services produced by workers, it would have to be carried out through an expenditure of income. In this case, however, the source of income would remain obscure and it would be logically impossible to explain the very existence of monetary economies. Indeed, it is only through an absolute exchange between money and output, where money is a mere numerical form and output a quantity of heterogeneous physical goods, that physical goods acquire a numerical form and money is transformed into income. As argued by Schmitt (1998–99a), the only absolute exchange giving rise to income coincides with the payment of wages, the only transaction allowing for the financial deposit of output on the assets side of banks' balance sheet or ledgers.

Only *individuals* are credited and debited. Neither land, nor capital can be the subject of credits and debits. [...] Once again, it is inconceivable for a capital to be credited or debited or 'credited-debited', only the 'human factor' can. It directly follows that human labour is the sole factor of production, provided production is understood in

its only scientific sense: an activity by which incomes are added to the riches of a national economy.

(*ibid.*: 52, our translation)

Now, the fact that capital is no macroeconomic factor of production does not entail that interest is no macroeconomic income at all. This is precisely the challenge facing any economist attempting to provide a satisfactory explanation of interest: to prove that, although capital does not add any positive value to produced output, interest is a macroeconomic income. Let us be as clear as possible. Fixed capital is beyond any doubt a determinant factor with respect to physical productivity. Thanks to fixed capital accumulation, the number (and sometimes the quality) of produced goods has significantly increased all over the world since the institution of banks. It is a fact that the increase in physical productivity has led to a corresponding increase in use values. The question that has to be answered, however, does not concern use values but exchange values, and in this respect things are slightly more complex. As we know from Schmitt's quantum macroeconomic analysis, what the Classics called exchange value is the result of the absolute exchange between money and output occurring at the moment of the payment of wages. Exchange values, therefore, do not vary according to physical productivity, which cannot be considered as a direct cause of income precisely because income is the macroeconomic result of this merging of form (money) and content (physical output). Hence, at this point it would seem necessary to conclude not only that wages are the sole macroeconomic factor of production, but also that their amount defines the totality of income formed in any economic system. If this were all there is to say about this topic, we would have to infer that interest does not exist as a macroeconomic income, its nature being merely that of a microeconomic income derived from wages.

As a matter of fact and as clearly intuited by some great economists of the past, interest does not have its origin in a mere redistribution of wages. Like wages, its nature is macroeconomic, but, unlike wages, it is not the result of the payment of a macroeconomic factor of production. Since the payment of wages is the only direct source of macroeconomic income, interest has to be derived from wages and has nevertheless to be added to them. How can these two requirements be fulfilled without running into a contradiction? The answer to this question passes through a correct understanding of the true nature of fixed capital and of its relationship with time.

As shown in Chapter 7, in its simplest expression capital is the form that enables income to subsist in time. As soon as it is created by production, income is transformed into capital and it is in this form that it can cross the interval of time separating its creation from its final destruction (consumption). This first form of capital is called capital-time by Schmitt (1984a) for a twofold reason: first, because income is but a quantum of time and, secondly, because capital-time owes its existence to the passing of time.

It is in its second form, that is, as fixed capital, that capital acquires a macroeconomic dimension. Whereas capital-time is just the momentary transformation of an income that is bound to recover its initial form to be definitively spent and destroyed through its final expenditure, fixed capital owes its existence to a transaction, to wit investment, which defines a macroeconomic saving, that is, the irreversible transformation of income into capital. While the transformation of income into capital-time is reversible, its transformation into fixed capital is not, because it takes place through the investment of a new production of instrumental goods.

To be exact, even capital-time itself is the result of an investment. When wages are paid to wage earners, they are immediately lent by banks, where they are formed as deposits, to firms, which, at the same instant, invest them in the purchase of current output. However, this 'intermediate' purchase, which transforms income (wages) into capital-time, does not enable firms to become the final owners of current output. What they spend is an income owned by wage earners and lent by banks, not their own income. The initial investment of firms made possible by the loan of wage earners' deposits with banks defines the formation of a stock and does not entail the destruction of the income saved by wage earners and lent by banks. This is so much so that when income holders decide to spend their earnings for the final purchase of output (formed as a stock with firms), capital-time is immediately re-transformed into income. Firms' initial debt is cancelled through the sale of their stock and so is capital-time.

What happens with the formation of fixed capital is a significantly different story. This time the income invested is irreversibly lost to consumption and defines a macroeconomic saving. It is through the investment of profit that fixed capital is formed and the expenditure of profit in the financing of a production of instrumental goods is final so that firms become the owners of fixed capital goods. Whether fixed capital formation occurs in an orderly or in a pathological way, the investment of profit is a macroeconomic transaction that gives

rise to an equivalent macroeconomic saving. It is precisely this transaction that verifies Keynes's fundamental identity between S and I . Macroeconomic saving is always necessarily equal to macroeconomic investment, since it is through the investment of profit and only through it that part of currently produced income can be definitively subtracted from consumption.

The relationship between capital and saving is that of a perfect reciprocity in the exact sense that capital is always formed through saving. Capital-time derives from the saving of wages. Indeed, the payment of wages defines the formation of a bank deposit to the benefit of wage earners. Now, the fact that wage earners own a claim on the bank deposit they have been credited with means precisely that they have saved their income and lent it to their bank. The following step is automatic and occurs at the same instant: consistently with the principle of double-entry bookkeeping, the income deposited on the bank's liabilities side is lent to firms and transformed into capital-time through its reversible investment in the formation of a stock of wage goods. All the different analytical steps described here are simultaneous and imposed by the use of bank money, which is an additional proof that capital owes its existence to the institution of banks and to their compliance with the rule of double-entry bookkeeping.

The reciprocity between saving and capital is also present in the case of fixed capital. This time the income saved by households and invested by firms takes the form of profit. Through the expenditure of wages for the final purchase of wage goods, consumers transfer part of their income to firms, which spend it to finance a production of investment goods. The income transferred by consumers to firms is, by definition, an income that is not consumed, that is, a positive saving. If part of this saving formed as profit were redistributed by firms as dividend or interest, it would define a capital-time, while the totality of what is invested in the production of instrumental goods gives rise to fixed capital. It is therefore in the form of fixed capital that macroeconomic saving is conserved through time.

As we have already observed, the main difference between capital-time and fixed capital lies in the reversibility of saving characterizing the former with respect to the irreversibility of the latter. The income saved and transformed into capital-time at the moment of its very formation (the payment of wages) is dis-saved, and retakes its initial form when income holders spend it for the final purchase of wage goods. Hence, capital-time exists as long as income is saved, the reversibility of saving being the mark of what we can call microeconomic saving.

In this respect, capital-time is also microeconomic, its existence being only temporary and its result the building of a bridge between present and future (Keynes 1936/1946: 293). On the contrary, fixed capital will never regain the form of income. What firms invest in a new production is forever transformed into capital, and, since what is invested is an income saved by households, it immediately follows that the formation of fixed capital defines that of a macroeconomic saving. In other words, fixed capital is the final form taken by income once the latter is irreversibly saved.

Interest is undoubtedly paid both on capital-time and on fixed capital, but on what ground can its existence be justified? Could interest be justified if only capital-time existed? As surprising as this might appear, the answer is no, because in this case no additional macroeconomic income would be formed in the system. Every income is generated, directly or indirectly, by production, and capital-time does not entail any additional production. Nevertheless, could it not be said that the presence of capital-time justifies that of interest even if only as a microeconomic income? The traditional argument runs as follows. Since capital-time allows to bridge present and future, it makes it possible for income earners to postpone their expenditure, thus enabling other households to borrow their savings and spend them in their stead. A positive interest is required to encourage income earners to save and represents a compensation for their effort to avoid spending their earnings at once. But how to explain that lenders are indeed those who make a sacrifice and that it is thus perfectly understandable that borrowers are asked to give up part of their future income in order to compensate them? Is it acceptable to answer, as done by some economists of the past, that individuals tend to overvalue present with respect to future goods and this is why they have to be encouraged to save by offering them a positive interest? Psychological considerations of this kind are highly hypothetical and subject to continuous variation. One could maintain equally that lenders should pay interest to borrowers on the ground that, because of technical progress, when the loan will be paid back lenders will be able to purchase more and better goods than those purchased by borrowers at the moment the loan was granted. According to this thesis, the sacrifice would fall on borrowers, who would have to be compensated by lenders. Finally, in the absence of a macroeconomic interest derived from fixed capital, no true scientific explanation of microeconomic interest can be offered.

Our previous conclusion is in line with that claimed by Böhm-Bawerk and Fetter, namely that the determination of economic interest

is logically prior to that of interest on contract loans. Since its formation, capital-time can be transformed back into income and lent by banks to consumers. In this case the loan granted by banks is financed by the income initially saved by income earners, who are entitled to be paid a positive interest. Banks play the role of intermediaries: they are paid a positive interest by borrowers and they pay an interest to the lenders, owners of bank deposits. As previously done by some of the past greatest economists, the existence of this interest on contract loans has to be anchored in that of a more fundamental interest derived from production. It is from fixed capital that this fundamental or macroeconomic interest comes. When firms invest their profit in the production of fixed-capital goods, households lose forever part of their income, which is definitively transformed into fixed capital. As we already know, this means that fixed capital is the form eventually attributed to macroeconomic saving. It is therefore correct to say that a sacrifice is necessary for the formation of capital, a sacrifice that explains both why interest is the remuneration of capital and how it is derived from it. The key to the understanding of interest is the macroeconomic saving defined by fixed capital. The income lost to consumption gives rise to a new production, and instrumental or fixed-capital goods are the result of the second, identical face of macroeconomic saving: macroeconomic investment. Interest is a kind of compensation households are entitled to because of their renunciation to consume the totality of their income, and its source is to be found in the increase in labour's physical productivity brought about by the presence of instrumental goods made possible by the investment of the income saved by households.

Interest rates and economic and financial crises

Let us refer to the neoclassical concept of the natural rate of interest. A key element in Wicksell's (1898/1965) analysis of interest and prices, the natural rate of interest is the level the rate of interest would reach in an economy in a state of general equilibrium, where normal profit would be equal to interest. In Wicksell's definition, the natural rate of interest is entirely expressed in real terms, that is, in real goods, and money is totally excluded. '[The natural rate of interest is] the rate of interest which would be determined by supply and demand if no use were made of money and all lending were effected in the form of real capital goods' (Wicksell 1898/1965: 102). Since our economies are of a monetary nature, Wicksell's analysis can be considered obsolete and

rejected altogether. Yet, it is also possible to reinterpret his concept of the natural rate of interest consistently with the monetary theory of production. Indeed, Wicksell derives the natural rate of interest from the profit realized by entrepreneurs as 'the amount by which the total product (or its equivalent in other commodities) exceeds the sum of the wages, rents, etc., that have to be paid out' (ibid.: 103). It is therefore legitimate to give the concept proposed by Wicksell a more general meaning and to identify it with that of the natural rate of profit, that is, the ratio between the profit realized in a given economy at a given moment in time, and the fixed capital available in this same economy.

Let us now introduce another concept proposed by Wicksell (1898/1965), namely the money (or market) rate of interest, that is, the rate of interest on contract loans. This money or market rate of interest is determined on the financial market starting from the policy rate of interest decided by central banks *vis-à-vis* their counterparties (commercial banks as a general rule).

Rather than distinguishing between 'real' and 'money' rate of interest, as done by Wicksell and neoclassical economists, we consider both the natural and market rates of interest as monetary magnitudes and we investigate their mutual relationship. In particular, we want to know what happens when the market rate of interest is lower than the natural rate of interest. According to Wicksell, such a situation would allow entrepreneurs to realize a surplus profit, which would attract new entrepreneurs and, via an increase in prices, induce the formation of a new equilibrium. Our analysis is entirely different, since it rejects both the neoclassical dichotomy between 'real' and 'monetary' magnitudes, and the general equilibrium approach. However, it is interesting to observe that the result reached by quantum monetary analysis is not inconsistent with Wicksell's as long as the natural rate of interest is identified with the rate of profit of the economy taken as a whole. To avoid any possible confusion, let us abandon Wicksell's analysis and consider the macroeconomic rate of profit and its relationship with the market rate of interest.

The existence of a positive gap between the macroeconomic rate of profit and the market rate of interest has marked the evolution of our capitalist economies for several decades. Indeed, it is the very presence of this positive gap that has enabled economic growth and capital accumulation. As long as the general rate of profit is substantially higher than the market rate of interest, it is worth investing in the production of new instrumental goods, which entails an accumulation of fixed capital and an increase in physical productivity. The problem is to know

whether the process of fixed-capital accumulation can go on forever or it is bound to come to a halt, as intuited by Wicksell. One condition is crucial in this respect: capital has to be remunerated. In other words, a positive interest has to be paid to the owners of fixed capital, which, as shown in Chapter 7, in a pathological economic system, as it exists today, are the disembodied set of depersonalized firms. Now, interest is paid out of profit, which means that as more capital accumulates, more profit is needed to remunerate it. No problem would arise if profit grew at the same or at a greater pace than capital, which is unfortunately not the case. As shown by Schmitt (1998–99a), since profit is derived from wages, its growth is limited by the amount of wages: at most profit can be equal to wages, which means that the limit of expansion of profit is reached when the production of profit goods is equal to that of wage goods. This means that, in any given period, the amount of interest cannot exceed the limit imposed by the payment of wages, whereas any new production of instrumental goods leads to a corresponding increase in fixed capital accumulated. '[C]apital increases without interruption, period after period, as a consequence of every new investment while, in each period, the production of interest reaches a limit that it can never exceed' (Schmitt 1998–99a: 129, our translation).

As is well known, if in a ratio the numerator increases less than the denominator, the ratio tends to fall. This is precisely what happens to the macroeconomic rate of profit. As intuited by Marx, the limit imposed on the growth of profit and the constant increase of fixed capital leads to a tendency of the rate of profit to fall. 'It is the conjunction of this limit imposed to the production of interest related to an uninterrupted increase of the capital that has to be remunerated (in interest) that represents the terms of the objective or natural dynamics of capital, tendency of the rate of profit (interest) to fall' (Schmitt 1998–99a: 129–30, our translation). The more capital is accumulated, the more difficult it becomes to remunerate it, until a limit is reached beyond which capital accumulation has to be reduced, because otherwise profit will no longer be enough to finance the payment of interest.

Another way of showing the insurgence of a crisis due to capital accumulation is to refer the macroeconomic rate of profit to the market rate of interest. As time goes by and more fixed capital is accumulated, the macroeconomic rate of profit decreases and the gap between it and the market rate of interest tends to zero. When the two rates are aligned, a crisis is unavoidable, no further investment in the production of new fixed-capital goods being possible in the system.

It is true that a reopening of the positive gap between the rate of profit and the market rate of interest might be possible through a reduction of the latter. Yet, this possibility exists only insofar as the central bank of a given economy can reduce its policy rate of interest. Unfortunately, such a monetary policy has a limit, already reached or nearly reached by some of the most important central banks of the world at the time of writing, which is represented by the impossibility of generalizing the introduction of negative rates of interest. In the actual state of affairs, the margin for a further growth of capital accumulation in the most industrialized countries is extremely small if not inexistent. As a consequence, investment is reduced, deflation sets in, and unemployment grows.

The pathological process of capital formation and accumulation analysed in Chapter 7 is the cause of this constant reduction of the gap between the macroeconomic rate of profit and the market rate of interest, and of the ensuing economic crisis. It is also at the origin of a pathological capital that feeds speculation and increases the intensity of financial crises.

As we explained in Chapter 7, to date, the expenditure of profits for the production of investment and amortization goods gives rise to new bank deposits, as a result of the compensation paid to the relevant wage earners. Banks are thereby induced to provide further credit lines to any kinds of borrowers, because they have to pay an interest (even though rates of interest on sight deposits are close to zero in a number of countries at the time of writing) on the sum total of bank deposits. This is a monetary–structural factor of increasing financial fragility, which may lead to a crisis eventually: both banks and their debtors are led to inflate a credit bubble, the former in order to earn an interest compensating their depositors as well as their top managers and shareholders, while the latter increase their debt in order to keep or even to increase their standard of living (as consumers) or to capture financial rents across the global economy (as speculators).

The pathological capital resulting from the production and amortization of fixed capital indeed feeds speculation on domestic and foreign financial markets, where speculators – including banks and non-bank financial institutions – dispose of this capital looking for interest income that is not and cannot be enough to remunerate their expectations, since the total amount of income produced across the world economy is not enough to satisfy them altogether. Suffice here to recall that gross world product (about 72 trillion US dollars in 2012) is only around 10 per cent of the notional value of derivative instruments (700 trillion

US dollars as of June 2011) exchanged over the counter. Even considering the errors and omissions position in this kind of statistics, one notices that world production does not generate enough income to satisfy all income claims circulating across the financial market worldwide. When the majority of speculators' mood is positive, much of these claims are not exerted, so that the global financial situation does not represent a problem apparently. However, as Minsky famously argued (see Chapter 3), 'financial stability is destabilizing', as financial institutions are led to increase their speculative lending, and this, sooner or later, ends up in a financial crash. The ongoing increase in pathological capital resulting from fixed capital accumulation and amortization feeds this building up of financial imbalances, and aggravates the consequences of a financial crisis, as they in fact affect wage earners (via an increase in unemployment) beyond speculators and their creditors within the financial system of any country.

These domestic monetary–structural factors of financial crises are further reinforced by a structural disorder that denatures bank money in international payments, an issue that we address in the next chapter.

9

The International Dimension of Financial Crises

Is the present system of international payments an orderly one? And if not, why and what are the consequences? These are the questions asked in this chapter, which is devoted to the analysis of international transactions and their impact on financial crises. The inconsistency between the way money should be used in transnational payments and how payments are actually carried out between nations is not difficult to discern or detect: what should be a mere numerical means of payment becomes an asset that is exchanged against other assets. Yet money does not change its nature when it is used internationally, which is why transnational payments that do not comply with the vehicular nature of money are hopelessly pathological and the source of monetary disorder that feeds speculation and propels financial crises.

We will start by showing what is wrong with the system of international payments as it actually works. In particular, we will consider the problem of how money intervenes in the payments between countries and how these payments allow for the final settlement of transnational transactions. This analysis will lead us to investigate the impact of international payments on exchange rates and on the formation of a financial capital that is not backed by any production.

In the second and third sections of the chapter we will introduce two strictly related analyses that have led Schmitt (2012, 2014) to the discovery of the pathological duplication of countries' external debts. The first analysis concerns the problem of indebted countries' external debt servicing and shows how the payment of net interest on a country's external debt has actually a total cost of *twice* the amount of the interest due to foreign creditors. Resulting from a long lasting research started in the 1970s and completed at the beginning of the third millennium, the *theorem of interest* proves rigorously that in the present non-system of

international payments a mechanism of a purely macroeconomic nature duplicates the charge of the payment of net interest, with the indebted country suffering from a loss of real resources and, additionally, from a reduction in its official reserves.

The second analysis in some way encompasses the third, and shows that the pathological duplication induced by transnational payments entails the very formation of countries' external debts and not only the payment of net interest on these debts. The problem at stake here is what has been called the sovereign debt crisis. Wrongly identified with the debt of the public sector, a country's sovereign debt results from a pathological duplication affecting the payment of its net foreign expenditures, that is, the difference between its total imports (commercial and financial) and its total exports. By referring to Schmitt's (2012, 2014) groundbreaking investigation, we will introduce the reader to his discovery of the mechanism that multiplies by two, to wit, doubles the payment of countries' deficit any time it is financed by an external loan. In a nutshell, it will be shown that the payment of net total imports has a cost twice as high as the amount of the deficit itself and that this additional charge is the only cause of countries' sovereign debt.

The lack of a true system of international payments and its consequences

The present system of international payments has its origin in the gold exchange standard adopted at the Bretton Woods conference (1944) and whose basis were laid down at the Genoa conference held in 1922. Since its adoption, the gold exchange standard has undergone various changes, the most important ones being the 'suspension' of the dollar's official convertibility into gold and the passage to a multi-currency or reserve-currencies standard. However, despite these notable changes the system has substantially remained the same, with one or more national currencies being used to settle international payments. The currencies chosen as international standards are supposed to have an intrinsic positive value and their exchange rates are identified with their relative prices. Much as if they were goods, currencies are bought and sold and are believed to have a price resulting from their relative exchange and depending on their supply and demand.

The misconception of bank money extends far beyond national boundaries and its negative consequences are particularly striking at the international level. Let us highlight some of them.

Issued by banks, money is a double entry in their books and is first of all defined as a spontaneous acknowledgment of debt. It is this acknowledgment of debt that is used as a means of payment between economic agents. What is crucial to understand here is that, while they are carried out or 'conveyed' by means of this acknowledgment of debt, national payments are not settled in money. By its very definition an IOU is just a promise to pay and it would be nonsense to claim that any economic agent or institution is entitled to pay by acknowledging her/his or its debt to the creditor. In order to pay and thus get rid of her/his debt, the debtor has to give up a positive amount of income and there can be no doubt whatsoever that income cannot be created *ex nihilo* but must result from production. This is indeed what happens within any national payment system, where bank money is given a real content by current production and where payments are settled, *through* money, *in* income. Unfortunately things change when a national currency is used to convey international payments. This time the acknowledgment of debt of a given country is not a simple means but the very object of transnational payments. The debtor country considers itself paid as soon as it is credited with a given amount of the reserve-currency chosen as international standard, which is tantamount to saying that it accepts to be paid, that is, to free its debtor from any engagement in exchange of the debtor's remittance of its own IOU or of the IOU of a third country.

The logical principle according to which nobody pays by getting indebted is simply ignored in the field of international payments, which would better be called a system of non-payments or a non-system of international payments. Now a question arises that requires a deeper analysis of the way payments are carried out between nations, namely how can countries be credited with a sum of foreign currency given that monies are but flows? As we have seen in Chapter 1, money is issued by banks in an instantaneous circular flow that prevents it from being stocked anywhere in the real world. This is so because double-entry bookkeeping requires the simultaneous debit and credit of the payer and credit-debit of the payee. Whether this principle is acknowledged or not, whether it is complied with or not, it imposes itself. In no conceivable circumstance can it happen that money does not immediately flow back to its issuing bank. What the payee obtains in exchange for her/his sales is therefore not an amount of money, but a deposit formed with the bank of the payer. Through the instantaneous flow of money, the payee obtains the ownership over a bank deposit whose object is a sum of income. This clearly means that a

payment between two economic agents residing in the same country can be carried out only if a positive income has been formed in the economy as a whole. When a payment occurs between two residents in different countries, what has to be determined is whether the country to which the payment is addressed receives, as settlement of its commercial or financial exports, part of the importing country's income or not.

The answer to this question is more complex than it might appear, since it is at the same time positive and negative. If country A pays for its net commercial imports from country B by crediting B's banks with an amount of money A, MA, double-entry bookkeeping guarantees the immediate reflux of MA into A's banks. This means that what is obtained by B is a claim on bank deposits formed with A's banks, that is, a claim on part of country A's income. Yet, it is also true that not a single unit of A's income is transferred to B. The income formed in country A remains deposited with A's banks and what B obtains is merely a claim on A's bank deposits. If, instead of occurring between residents in different monetary areas, the payment were carried out between residents of the same monetary area no problem would arise, since bank deposits would be denominated in the same currency. This is not the case here, for countries A and B use each its own domestic currency. In such a framework two options are conceivable: either the system of international payments is consistent with the vehicular use of the money chosen as means of payment or the money in which payments are carried out is considered as an asset with a liberating power. In other words, either payments are *conveyed* by money or they are *settled* in money. At present, it is a matter of fact, international payments take place in line with the second option.

Despite the vehicular nature of bank money, a sum of money A is entered on the assets side of B's banks ledger as a result of country A's payment. As clearly exposed by Rueff (1963), the fact that a positive sum of money A is entered as an asset by B's banks while the totality of A's income is still deposited with A's banking system proves that MA is the object of a duplication: it is simultaneously entered as an asset in B's banks and deposited in A's banks. 'Entering the credit system of the creditor country, but remaining in the debtor country, the claims representing the deficit are thus doubled' (Rueff 1963: 324). What is annoying with duplication is that the net exporting (or creditor) country considers itself paid with a sum of money that is nothing else than an empty duplicate of what is entered as a positive income in the banking system of the net importing country.

Let us repeat it once again: by its very nature, bank money is a flow so that as soon as country B is credited with a positive amount of money A it is also necessarily debited with the same amount. If the system of international payments were so structured as to explicitly comply with the vehicular nature of money, B's banks would never enter an amount of MA as a net asset. In exchange for its net commercial exports, country B would obtain an equivalent amount of financial claims over country A's present or future output and not a sum of nominal money, an empty duplicate of zero value. However, it is a well-known empirical fact that net commercial imports are settled in money. In our example this means that, even though money A flows instantaneously back to its point of departure (A's banks), it is simultaneously entered as a net asset into B's banks. It is thus correct to claim at the same time that B obtains the ownership over part of A's income and that it does not obtain it. Indeed, country B gives up its right over A's income at the very instant it retains a sum of money A as counterpart of its exports.

In his 1984 book on international payments, Schmitt (1984b) shows that Rueff's duplication is at the origin of an inflationary increase in the capital available in the net exporting country. 'Financial capital increases *twice* in the net exporting country; in national money and in foreign exchange. One of the two increases can only be *fictitious* since the national economy's gain *vis-à-vis* the rest of the world equals the amount of net commercial exports – not *twice* this amount' (Schmitt 1984b: 43, our translation). Indeed, on the basis of the amount of money A entered as an asset in its ledger, B's banking system issues an equivalent amount of money B. This emission is but the monetization of the external gain obtained by country B thanks to its net commercial exports and defines the legitimate capital formed in B as a result of its international transactions. The foreign exchange (a sum of money A) still entered on the assets side of B's official reserves (its necessary point of arrival) is the fictitious gain to which Schmitt refers in the passage quoted above. Deprived of any real content, it pathologically increases the financial capital of country B and is thereby the cause of an inflationary gap of international origin.

The process of duplication briefly described here is symptomatic of the anomalous use of one or more national currencies as an international settlement unit. Erroneously transformed into an object of exchange, money A is considered at par with, or equivalent to a real good whereas, in reality, it is a purely nominal unit that increases, pathologically, the international financial capital available offshore. Initially entered as an asset of B's commercial banks, money A is then transferred to B's central

bank, where it is entered as an asset in B's official reserves, and is finally invested on what used to be called the euro-market, that is, offshore financial centres where it increases the speculative capital that is unanimously considered as the main source of the on-going financial crisis at the time of writing.

Bank money is a dematerialized means of payment and, as such, it can neither be purchased nor sold. In its circular flow it conveys real goods, present and future, from the payer to the payee and vice versa. This implies that, in a system respecting its vehicular nature, money would never be the object of any payment. Real goods, financial claims included, are the only 'objects' of exchange, not money, which is but a numerical means of exchange, 'the great wheel of circulation' (Smith 1776/1991: 256). Yet, if a system of payments like the one adopted internationally up to now fails to acknowledge the very essence of bank money and identifies it with a net asset, the set of real goods (commercial and financial) is artificially increased by the inclusion of money. From a simple means, money is transformed into an object of exchange and as such it is supplied and demanded for its own sake. Comparable to a real good, this 'denatured' money has its own market, where it can be bought and sold, and the rate of exchange between national currencies is identified to their relative price.

Real goods are purchased and sold and money should simply be the (numerical) means through which these transactions are conveyed. This is not what happens today: money is itself an object of exchange whose price varies according to the interplay of supply and demand. On the foreign exchange market, national currencies are supplied and demanded as if they were real goods, and their prices fluctuate according to these interacting 'forces'. This state of affairs acquires all its dismal aspect once it is remembered that the market where currencies are bought and sold is fed by huge amounts of financial capital pathologically formed through duplication. Quantum analysis confirms what an increasing number of bankers and financial experts have maintained for a long time: speculation is by far the main cause of exchange rate erratic fluctuations. What has still to be fully understood is that the source of speculation is not to be identified in the existence of speculators, but in the growing availability of speculative capital. If no pathological capital were available, speculation could not be fuelled and would not have the disruptive effects that our economies actually experience.

The crucial role played by speculation in the building up of financial crises does not need to be emphasized. What has to be stressed instead is that speculation is made possible by a pathological mechanism leading

to the formation of an inflationary, financial capital dissociated from real output. In Chapter 7, we have seen how such a capital is formed within a national economy; in the present chapter we are on the verge of verifying that the actual system of international payments is also at the origin of a financial capital of an entirely pathological nature. The wrong conception of money on which this system is built is responsible for a nominal duplication that pathologically increases the amount of financial capital. It is symptomatic that the expression coined to denote this pathological capital is that of *financial bubble*, which clearly conveys the idea of a nominal or empty excrescence whose presence has an unsettling impact over the financial and the economic sectors both nationally and internationally. As far as international transactions are concerned, the financial bubble increases dramatically the level of uncertainty and erratic fluctuations affecting the foreign exchange and the stock markets. It is because of the financial bubble that speculation has so much developed as to represent by far the main force acting on these markets. In this respect let us simply observe that, according to the *International Monetary Fund Currency Composition of Official Foreign Exchange Reserves*, in 2013 the official reserves available worldwide were about 11.6 trillion dollars, while the amount of *daily* transactions on the foreign exchange market were of 5.3 trillion dollars according to the *Bank for International Settlements Triennial Central Bank Survey*. These figures show that all central banks taken together could not counter-balance the weight of speculation for more than a couple of days. This is a clear sign of the overwhelming expansion of the financial bubble in today's non-system of international payments.

The final payment of net commercial imports *in money* (and not simply *by means of money*) is at the origin of a duplication of financial capital that is one of the main sources of the financial bubble's formation and growth. Another way of proving Rueff's intuition about the pathological process of duplication characterizing the present (non-)system of international payments has been proposed by Schmitt already back in 1984 and relates to the formation and servicing of countries' external debts.

The double charge of external debt servicing

Let us consider the case of an indebted country, A, facing the rest of the world, country R, and paying 10 units of money R, MR, in billion dollars, as net interest on its accumulated external debt. Let us also assume that 1 unit of money A, MA, exchanges for 1 unit of MR, $1 MA = 1 MR$, and

that country A's total imports (net interest excluded), IM , are equal to its total exports, EX :

$$IM \equiv EX \quad (1)$$

where IM represents A's total expenditures (commercial and financial) and EX stands for its total receipts, both of commercial and financial origin. The relation $IM \equiv EX$ is known as the balance-of-payments identity.

The payment of net interest, in , on country A's external debt increases A's expenditures by an amount of 10 MR and seems to overthrow identity (1). Yet, the payment of in must take place in MR and country A has already spent the totality of its resources in MR, obtained through its foreign sales, EX , in order to pay for its imports, IM . This means that A has to borrow 10 MR to pay for in , so that the balance of payments identity takes the following form:

$$IM + in \equiv EX + LD_{in} \quad (2)$$

where LD_{in} stands for loan disbursements (the new external loan granted by R to A) of an amount equal to in (10 MR). Hence, apparently nothing is wrong with the payment of net interest, the increase in country A's expenditures being perfectly matched by an equivalent increase in its receipts. Appearances, however, can be deceptive and lead to wrong conclusions. This is indeed the case of the payment of in .

As clearly stated by Schmitt (2000, 2005, 2007, 2012), interest on foreign loans defines that part of the yields derived from the investment of the capital borrowed abroad that is due to foreign lenders. A part of A's domestic resources are due to R, because interest establishes the ownership of R's residents (the lenders) over a production of A's economy equal to in . In our numerical example, a product of 10 MA (= 10 MR) is owned by R and has therefore to be transferred cost free from A to R. This is the meaning of the expression 'unrequited transfer' used by the International Monetary Fund to illustrate the nature of interest, and this explains why the payment of interest is entered into the current account of the balance of payments. Being an amount of domestic resources that is transferred to foreign creditors, interest is necessarily part of the indebted country's exports, which means that it is necessarily financed by the current account of the latter country.

The consequence of in being part of country A's exports is that an equivalent amount of A's exports is obtained by R at zero cost, without any real counterpart. Now, the non-payment by R of part of its imports

reduces by in the amount of A's receipts. If we denote by ex_{in} the amount of country A's exports that are transferred to R, we immediately realize that A's total receipts fall short of its total expenditures:

$$IM + in > EX - ex_{in} + LD_{in} \quad (3)$$

Inequality (3) results from the reduction in A's receipts brought about by the fact that a part equal to in of A's exports is not paid by R. The equality required by the need to respect the balance of payments identity is restored through a decrease, equal to in , of A's official reserves.

$$IM + in \equiv EX - ex_{in} + R_{in} + LD_{in} \quad (4)$$

where R_{in} represents the amount of its official reserves that country A has to sacrifice to bring its total expenditures, right hand side of (4), to level with its total receipts, left hand side of (4). As a result, the cost for country A of the payment of its net interest on external debt is multiplied by 2: the loss, equal to in , in official reserves adds up to the new external debt generated by LD_{in} , also equal to in , so that the final charge is equal to $2in$. To cover for the net interest on its external debt, equal to ten MR, country A and its residents have to endure a total loss of 20 MR.

Let us follow Schmitt (2005) and distinguish between the monetary and the real payment of net interest. As well established, both in theory and facts, the payment of in in real terms requires A to give up for free what R already owns. A product valued ten MA of A's domestic economy becomes the property of R from the moment in falls due. The unrequited transfer to R of part of A's domestic product is conveyed through A's exports, a part of which, equal to ex_{in} , is therefore not paid by R. The difference, in money terms, between A's total expenditures, $IM + in$, and A's total receipts, $EX - ex_{in}$, is thus no longer equal to in (ten MR) but to $2in$ (20 MR). Country A's new loan, LD_{in} , covers half the difference, the other half being covered by A's official reserves. An increase in its external debt, equal to ten MR, and a decrease in its official reserves, also equal to ten MR, are the two, additional charges of A's payment of net interest on its debt. On the whole, it costs country A 20 MR to pay for its net interest of ten MR.

Two payments of in add up to one another, one *real* and the other *monetary*. The payment in real terms is the unrequited transfer of A's domestic resources, the monetary payment is the additional amount of ten MR paid to A's creditors. The problem arises because the real payment of in entails the non-payment of ex_{in} by R. Indeed, it is A itself

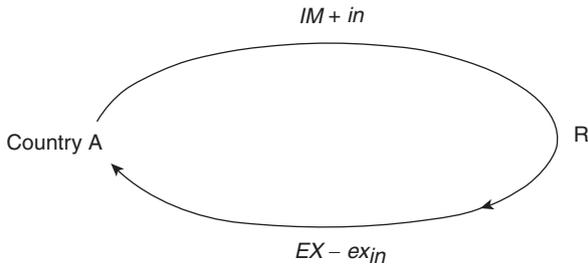


Figure 9.1 The double charge of the payment of in

that, by paying in to R, gives the rest of the world the sum of money income necessary to purchase part of its imports. Since a part equal to in of A's real exports is already owned by R, it would be inconsistent to ask R to pay for its purchase. The transfer to R of what R owns implies that ex_{in} is finally paid by A. As a consequence, the payment of in reduces A's receipts from EX to $EX - ex_{in}$. On one side, A's total expenditures are increased from IM to $IM + in$, on the other side, its total receipts are decreased from EX to $EX - ex_{in}$. Figure 9.1 describes this double effect clearly.

As shown in Figure 9.1, the payment of in by A finances an equivalent part of R's imports, thus reducing by ex_{in} the sum of MR obtained by A as payment of its exports EX . Country A and its residents have therefore to cover *two* deficits: one due to the *real* payment of in , which reduces A's receipts by ten MR, and one due to the *monetary* payment of in , which increases A's expenditures by ten MR, from IM to $IM + in$. Both deficits need to be financed and this brings to 20 MR the amount that country A has to find to restore equilibrium. Ten MR are found through a new loan, while the other ten MR are taken out of A's official reserves. Since a reduction in reserves is a reduction in the country's credits on the rest of the world, the second charge of the payment of in is *de facto* an increase of A's external debits. On the whole, A's external debt is thus increased *twice* following its payment of in : once, because of the new loan incurred by A, LD_{in} and the other, because of the loss in its official reserves, R_{in} .

The pathological duplication of A's external debt can be derived from the double charge of the payment of in , but it can also be established directly, through the analysis of its very formation. This is what has been done by Schmitt (2012, 2014) in his last papers devoted to a critical investigation of the anomalous working of the current non-system of international payments.

The sovereign debt problem: Its pathological nature

Let us start from the following, simple principle: a country gets indebted when it must recur to a foreign loan in order to finance its net disbursements. Hence, for example, if any given country A does not benefit from a positive amount of foreign direct or portfolio investments and if its trade balance is negative, it will necessarily have to borrow from the rest of the world, R, the sum needed to pay for its net commercial purchases. What really matters here is the difference between a country's expenditures and its receipts. When A's monetary outflows in favour of R are greater than its monetary inflows from R, the balance is restored through a foreign loan whose consequence is an increase in A's external debt. The main sources of monetary inflows for A are commercial exports, foreign direct and portfolio investments from R, and interest on its foreign credits (paid by R). Its main causes of monetary outflows are commercial imports, foreign direct and portfolio investments to R, and interest on its foreign debt (paid to R). Foreign loans are also a source of monetary inflows, when they are obtained from R, or outflows, when they are granted to R. Yet, what must retain our attention here is the fact that the sum borrowed abroad in order to balance a country's inflows with its outflows has an impact on its international investment position.

Nothing disturbing has been said so far. Indeed, no one will dispute the fact that, like any other economic agent, a country must balance its sales and its purchases and that it does so on the set of markets available internationally. To be precise, a country balances its sales and purchases on the commodity and financial markets taken together, while its residents do it on the set composed by the labour, commodity and financial markets. It is thus plain that if a country runs a deficit in its trade balance, it has to match this deficit with an equivalent financial credit and vice versa. Let us consider the simplest possible case, when country A balances all its foreign transactions except one: its trade on commodity markets. It is easy to observe that, in this case, it will balance its net commercial imports through an equivalent sale of bonds on the financial market, that is, by borrowing abroad the sum needed to equalize its foreign purchases and receipts. The expected result would be that if A borrows for say x billion dollars from R, its foreign debt increases by x billion dollars.

Schmitt's recent analysis of the way countries' external debts are formed leads to another, astonishing result: even though its external borrowing amounts to x billion dollars, A's external debt increases by twice this amount, that is, for a total of $2x$ billion dollars. This

conclusion is so unexpected and contrary to common sense that it sounds absurd. Is it not true that if any domestic resident of country A borrows a sum equivalent to y dollars, either within her/his own country or abroad, she/he incurs a debt equal to y ? Why should this not hold true even for countries? Are countries not an intellectual construct arrived at through the mere aggregation of their residents? Only apparently simple and straightforward, these questions call for a thoughtful answer and deserve our careful attention. Let us analyse them in the reverse order in which they have been laid down.

The relationship between a country and its residents is a close one, so much so that no country would exist in the total absence of residents. Yet, this first consideration is not enough to conclude that, from an economic viewpoint, a country's situation is but the exact image of that of its residents so that, at the end, it does not differ from the sum of its residents' situations. As a matter of fact, the concept best suited to define the relationship between a country and its residents is that of *set*. A country is a set, a whole, and its residents are the elements of this set. As for any other set, a country cannot be identified with the simple sum or aggregation of its elements. This means that, in economics, the concept of country is of a macroeconomic nature and has to be analysed accordingly.

Our definition of a country as the set of its residents might appear dogmatic to a superficial reader used to consider macroeconomic magnitudes as the result of aggregation of microeconomic variables. Yet, some simple considerations about the nature of national currencies suffice to show that countries do not simply reproduce the *status quo* of their residents. As unanimously recognized, in fact, money is banks' spontaneous acknowledgment of debt, the IOU of a specific banking system. This is to say that any transnational payment carried out by a country's residents in their national currency transfers abroad an IOU of their national banking system, that is, an acknowledgment of debt of the country itself. The meaning conveyed by the expression 'national currency' is clear; it denotes the currency of a nation, its own IOU. Now, the use of a national currency enables residents to settle their transnational transactions, to pay for their purchases abroad and thus get finally rid of any debt incurred to foreign residents. Once all these payments have been carried out, no resident runs an external debt any longer and if countries were identified with their residents we would have to conclude that the country itself is no longer involved with the rest of the world. In reality, this is not what happens, since the final payment made by its residents gives the rest of the world a credit over the country itself,

as represented by its national banking system. As this example clearly shows, a country can be indebted even if no one of its residents is. Analogously, it is enough to consider the opposite situation to verify that a country can run a credit even though all its residents have been finally paid and none of them is still a creditor with respect to R. However, the question that has to be asked in this respect is whether it is correct to claim that a country running a deficit, that is, whose total expenditures exceed its total receipts, incurs a positive external debt even though its residents have paid the totality of their foreign purchases, net purchases included. Before thoroughly analysing this question, let us elaborate on the theoretical and practical framework in which it arises.

The economic existence of nations is unavoidably linked to that of national currencies. If countries were to abandon their monetary sovereignty, they would lose their specific economic existence and become regions of a unique, multicultural nation. Yet, the transformation of countries into regions is not a simple matter of replacing their national currencies with a common currency chosen as a standard. A true process of monetary unification requires the implementation of a unique banking system enabling the final payment of inter-bank transactions across national borders. As long as payments between residents and non-residents fail to be incorporated into a system of real-time gross settlements headed by a common central bank and national currencies replaced by a unique currency issued by all the banks operating on the territory, countries will continue to exist as economic entities, and their existence will require transnational payments to be settled accordingly.

The answer to our second question is a direct consequence of the definition of countries as sets. If countries were merely representing the sum of their residents, their indebtedness would have no proper existence and would not call for any specific analysis. Things not being so, it would be mistaken to infer that what is true for the residents must also be always true for their country. In particular, the fact that a country is the set of its residents and not their sum is enough to reject the claim that a country's external debt is nothing but the external debt incurred by its residents. It is up to monetary macroeconomics to investigate the economic nature of this existence in order to determine the involvement of countries in the transactions carried out by their residents and establish the rules a system of international payments must comply with to avoid the pathological addition of the macroeconomic payment (concerning countries) and the microeconomic payment (carried out by their residents).

If we now go back to our initial question, we can see that, although it is undoubtedly true that when the residents borrow γ their debt increases by γ , if the loan is obtained from abroad the country is also involved in the transaction. It cannot therefore be excluded, *a priori*, that the debt incurred by the country might end up being additional to the debt incurred by its residents. If this were to be confirmed, we would have the proof that the present system of international payments is pathological and that a reform is urgently called for to solve what is known as the sovereign debt crisis.

Before presenting Schmitt's analysis of the pathology affecting the formation of sovereign debts, let us clarify the meaning of the concept of sovereign debt. A country's sovereign debt being frequently identified with the debt of the State, it is first of all necessary to disprove this claim. In fact, once it is agreed that a country is the set of its residents, it becomes clear that a country's debt derives from the debt incurred by all its residents and not from only one of them, the State, irrespective of how important it may be. Furthermore, it should also be clear that while the State or any other resident can get indebted either internally or domestically (to other residents), or externally (to non-residents), a country's debt is necessarily a foreign debt. These simple considerations lead us to the conclusion that a country's sovereign debt is but its external debt and that it relates to the sums borrowed abroad by both its public and private sectors in order to cover for the difference between its total expenditures or imports and its total receipts or exports.

Now, if countries' sovereign debts were simply the image of the foreign debts incurred by all their residents (State included), the present sovereign debt crisis would not be pathological. It would simply be explained by the excessive external borrowing of the public and private sectors and would call for a set of measures apt to restrain public and private deficits. This is what the majority of economists firmly believes and what most governments are implementing, thereby imposing great sacrifices on their constituents and hampering an economic recovery systematically postponed and increasingly out of reach. In reality, if the sovereign debt crisis is so devastating it is because it is not due to countries' residents living beyond their 'means', but to a pathological mechanism inherent in the present system of international payments that gives rise to a debt for the countries themselves, which adds up to that of their residents. Thus understood, the sovereign debt is an anomaly whose cost lays heavily on countries, which are deprived of an increasing part of their internal resources to the benefit of what has

been called the financial bubble. Let us show how this happens and how greatly countries suffer from this anomaly.

Foreign surplus imports are paid both micro- and macro-economically

The central point of our investigation is the distinction between micro- and macroeconomic payments. It is essential to observe once again that if only microeconomic payments were involved no pathology would arise. Yet, such a situation could be envisioned only if regions of a single world nation replaced countries. This not being the case so far nor for the foreseeable future, the analysis of the present system of international payments has to choose between two alternatives: (1) macroeconomic payments, that is, the transnational payments of countries, are always perfectly balanced or reciprocal so that they always cancel out; (2) macroeconomic payments do not necessarily balance one another, in which case they add up pathologically to the microeconomic payments carried out by countries' residents.

Let us go back to our initial case, where a given country, A, borrows abroad the sum of foreign currencies needed to pay for its net commercial imports (total expenditures). As far as A's residents are concerned, in order to pay for their foreign purchases they simply must dispose of the required amount of national income. Once their domestic banks have debited them the sum due, A's residents are no longer indebted to their foreign counterpart, the exporters of goods and services residents of R, the rest of the world. Even though the microeconomic payment of A's net commercial imports has been successfully carried out by its residents, A is nevertheless forced to respect its balance of payments requirement, notably, it has to balance its trade deficit with an equivalent surplus of its capital and financial account. Unless one were prepared to claim that balance of payments are simply statistical constructs with no real impact on countries, one would have to acknowledge the existence of a problem that goes beyond the relationship concerning residents and non-residents.

An attempt to solve the conundrum could be the following. The payment carried out by A's residents (or, more precisely, by their domestic banks) conveys to the banks of R an acknowledgment of debt of A's banking system. Paid by the banks of A, the banks of R become their creditors: they become the owners of claims on the deposits formed in A's banks. We thus may have the impression that, as a consequence of the microeconomic payment carried out by its residents, country A balances its net commercial imports with an equivalent export of financial

claims, as required by the balance of payments principle. If this were correct, we would conclude that the implication of country A simply mirrors the payment of its residents so that no proper macroeconomic problem arises. However, a simple consideration is enough to realize that things are more complex and that the macroeconomic problem is still far from having been solved. Indeed, if the payment of A's residents were all that is required to settle A's net commercial imports, we would have to claim that the sale of financial claims on A's bank deposits through which country A balances its net commercial purchases enables A to borrow from R *a sum of its own domestic resources*. This conclusion is highly unsatisfactory, since what country A needs to obtain from its external borrowing is a sum of foreign resources, and not a sum of its own national income.

Another consideration settles definitively the question: the claims on A's bank deposits do not define a right over A's future output and therefore cannot be assimilated to an amount of (future) real goods given by A in exchange for its commercial imports. In this sense, claims on bank deposits are not financial claims since the objects of these items are substantially different: an acknowledgement of debt of banks (bank deposits) and part of the country's future domestic resources (financial claims). We are thus led to the following result. The involvement of country A is not the simple, passive consequence of the payment carried out by its residents, but requires the borrowing of foreign resources to cover for the difference between total expenditures and total receipts.

The conclusion we have reached so far should not come as a surprise, for it simply reiterates the well-known principle that, in the absence of external gains, a country balances its net imports through foreign borrowing. It is therefore in line with general thinking to claim that net imports are a cause of countries' external or sovereign indebtedness. The novelty of Schmitt's analysis rests on an astonishing discovery: when they recur to foreign borrowing, countries get twice as indebted as they should be. Hence, if country A borrows, say, x billion dollars abroad in order to finance its net foreign purchases, it will unavoidably end up with a new total external debt of $2x$ billion dollars. Let us, from the outset, stress the fact that such a pathological duplication is entirely of a macroeconomic nature. It has nothing to do with the microeconomic payment of A's residents, nor can it be imputed to the foreign borrowing of A's public and private sectors. If A's residents borrow abroad a sum of x billion dollars, their external indebtedness increases by this sum and certainly not for twice as much. The problem arises because of a pathological mechanism that, for lack of a true system of international

payments, *generates a monetary deficit for the country as a whole*. Resulting from the difference between the country's total monetary outflows and inflows, this deficit calls for an immediate compensation, which occurs through an increase in A's external debt.

The duplication of indebted countries' external debts

Let us refer to the arguments proposed by Schmitt in his 2014 paper on the double charge of external debts and the formation of sovereign debts. What we have to consider are the monetary flows relating to the external loan of x billion dollars obtained by country A and to its expenditure for the purchase of an equivalent amount of imports. As far as the real flows are concerned, it seems possible to maintain that a reciprocal exchange occurs between the foreign products (commercial and financial) imported by A and the claims exported by it. Actual real goods exported by R and future real goods exported by A balance, and we thus get the impression that money passively adapts to this state of affairs and that its flow is also perfectly balanced. This would indeed be the case if international payments were carried out through a system respectful of the vehicular nature of money and if countries, as sets of their residents, were subjected to the strict principle of double-entry bookkeeping. Since a system of international payments based on the systematic respect of the law of the necessary equality between each country's sales and purchases has not yet seen the light of day, the neutrality of money is not guaranteed and its flows can be balanced only through a pathological variation of the country's indebtedness.

The principle on which the analysis rests is that of double-entry bookkeeping. As we have repeatedly claimed, correctly interpreted this principle establishes that no economic agent can be credited (debited) with any given amount of money without being simultaneously debited (credited) with the same amount. Double-entry is such that any time a bank credits (debits) an economic agent it necessarily debits (credits) her/him to the same extent. It is this principle that establishes once and for all the flow nature of bank money. Now, as any other logical law, this principle applies always, whether the system of payments complies with it or not. Money is a flow irrespective of whether it is considered and recognized as such or it is wrongly identified with a net asset. Within any national system of payments, double-entry bookkeeping is enough to guarantee in most cases the vehicular use of money, exceptions being bounded to the actual pathological process of capital formation and amortization (see Chapter 7). At the international level, however, this is far from being the case. Equilibrium between the real

products exchanged by A, the country running a deficit, and R, the rest of the world, is not enough to guarantee the existence of an international monetary circuit, so that country A has to purchase at a positive cost the foreign currency required for the settlement of its net purchases, already fully paid by its residents in real terms.

The previous argument can be presented in two slightly different ways. The simplest consists in observing that A's residents pay in money A, MA, the totality of their foreign purchases, so that in no circumstances should their country incur a debt to R. It is true that R's residents demand to be paid in money R, MR, and it is also certain that, in today's non-system of international payments, country A has to borrow abroad this sum of MR. Yet, it is highly illogical to pretend that a second payment, the reimbursement of the loan obtained by A, must add up to the payment already fully carried out by A's residents. In an orderly system, country A would never have to borrow a sum of MR equivalent to the amount of income in MA paid by its residents.

A more complete and complex way of presenting the argument is to say, following Schmitt (2014), that the addition of the monetary payment carried out by country A in MR to the payment in real terms carried out by its residents in MA gives rise to *two external debts*. The first debt, equal to A's surplus imports (expenditures) – x billion dollars in our example – is incurred in order to restore the amount of A's domestic income initially reduced because of the payment of A's residents for their net imports. Through this foreign loan, A gets paid by R for the real products that A's economy will produce and export in a future period.

Let us introduce here the same numerical example used by Schmitt in his 2014 paper. Suppose A's total purchases to be equal to 11 MR (billion dollars) and its total sales to be equal to ten MR, also in billion dollars. Let $1 \text{ MR} = 1 \text{ MA}$ be the exchange rate between A's national currency and the currency of the rest of the world. Since A's actual imports, 11 MR, are greater than its actual exports, ten MR, the equilibrium in real terms between A's and R's external transactions is obtained by adding the export of future goods worth 1 MR to A's actual exports. The equilibrium of A's and R's balances of payments in real terms is thus granted by the first external loan, whose result is to transfer to R the ownership over a part of A's future production equal to the domestic production of R exported in surplus in the initial period. Unfortunately, however, this is not the end of the story. As things are today, country A has still to pay R an amount of foreign currency equal to 11 MR. Indeed, R's residents require to be paid in MR and country A can oblige only if it borrows 1 MR from the rest of the world. This second debt incurred by

A has an amount of money R as its sole object, and is therefore entirely pathological.

Money, whether nationally or internationally, is a mere numerical means of exchange and payment. As such it should never be considered as a good or asset. If it is, it would be wrong to infer that its nature has changed. Money remains what it is, but the fact that this is not recognized by the present system of international payments has the disruptive consequence that countries in deficit have to incur a debt in order to obtain a numerical vehicle they should obtain free of cost. If the present system of international payments were an orderly one, the respect of the balance of payments identity would confirm to the law of the logical identity between each country's sales and purchases, and the vehicular use of money would automatically result. In each international transaction both A and R would be instantaneously debited-credited and credited-debited with the same amount of money (MA, MR or any other money chosen as a standard). On these conditions, A would never have to borrow any net sum of MR and its foreign debt would not even exist. Today things are profoundly different, and A has to incur an additional external debt, which multiplies by two the burden on the country and, as a consequence, on its domestic economy. If, in our example, an external debt of one billion dollars is justified by the surplus of A's imports, its multiplication by two has no rational *raison d'être* and is the clearest sign of the pathology affecting our non-system of international payments.

A statistical data analysis confirms the pathological duplication of countries' external debt

Let us consider three countries (Italy, Brazil, and Mexico) as actual cases of the phenomenon we have just described. Our choice of countries is somewhat arbitrary and the reader can easily transpose the analysis to any other country of her/his choice. If we have decided to consider the case of Italy, Brazil, and Mexico, it is because we want to show that the pathology affects both developed and less developed countries, members or non-members of a monetary area.

The entirely pathological nature of the sovereign debt crisis

As shown by Schmitt's (2014) analysis and supported by statistical evidence, external debts are duplicated by a mechanism that adds a debt (resulting from the necessity to cover, in foreign currency, the difference between a country's external purchases and sales) to the debt incurred by its residents. Now, correctly defined a sovereign debt is precisely the

debt of the country as the set of all its residents. This means that if, as is indeed the case, the residents pay the totality of their purchases, no external debt of the country itself should exist logically.

In this respect, it would be wrong to claim that the debt of any given country, A, is nothing but that of its residents. A's debt is not an identical copy of the debt incurred by its residents. The most incisive proof resides in the fact that the debt of country A that can be ascribed to the payment of its residents is the one that guarantees the reconstitution of the domestic income, in MA, which this very payment reduces. Once this first debt has been factored, another debt remains outstanding, because the deficit of A has still to be paid in MR.

Even a quick look at statistical coverage by the main world financial institutions and at the relevant economic literature suffices to verify that it is almost universally believed that most countries are indebted and that their indebtedness has to be imputed to their having lived beyond their financial possibilities. Schmitt's analysis shows, on the contrary, that their sovereign debt results from a pathological duplication, which originates in the absence of a consistent system of international payments. In other words, countries' sovereign debts are entirely unjustified and it is highly unreasonable to compel them to adopt austerity policies that impoverish their populations and hamper their economic recovery. The expression 'sovereign debt crisis' is indeed appropriate, since it emphasizes the dramatic worsening of a state of affairs that affects both debtor and creditor countries. The former countries because they are forced to serve a debt that should not even exist, the latter countries because their economic development is restrained by the continuous impoverishment of their partners and their risk of default.

The situation is indeed quite simple: countries whose residents have paid all their external purchases, surplus imports included, are forced to borrow abroad the foreign currency necessary to convey this payment to R. If we refer once again to country A, the anomaly jeopardizing the present system of international payments is such that A is forced to pay in MR what it has already fully paid in real terms. A's net purchases of R's real products is indeed perfectly balanced by the real income obtained by R from A's residents. It is true that this income is recovered by A's domestic economy, but this is the result of a foreign loan that gives R the ownership over a future product of economy A. In our example, A's total imports of R's real products, equal to 11 MR, are balanced by an export of real products actually produced by A, equal to ten MR, plus an export, equal to 1 MR, of real products that economy A will produce in the future. The only difficulty here is to understand that R acquires

Table 9.1 The equality between A's real imports and its real exports

Country A			
Exports		Imports	
Exports of current output	10 MR	Imports of current output	11 MR
Exports of future output	1 MR		
Total exports	11 MR	Total imports	11 MR

the ownership over this future product of A from the moment it lends 1 MR to A's economy. Even though R's residents will import and pay A's future product only in a subsequent period, R finances it since the period its residents grant a loan to A. The fact that A will be able to reimburse this loan in the future by exporting part of its real product to R confirms that this very product is, since the period the loan is actually granted, the property of R. It is therefore correct to say that country A's total imports of 11 MR are balanced by total exports also equal to 11 MR (Table 9.1).

If the equilibrium between country A's imports and exports provided the free-of-cost use of the vehicular currency needed to convey the real exchange between A and R, no duplication would occur and the cost of A's net imports would be a single one: the cession of A's future products necessary to bring its real exports to the level of its real imports. Since this is not actually the case, country A has to incur an additional debt in order to obtain the amount of MR it needs to pay its net imports in foreign currency. Country A's sovereign debt results directly from this additional foreign loan corresponding to its second payment, purely monetary, of its trade deficit.

The first debt incurred by A's residents in order to balance their net imports with an actualized export of future products is perfectly legitimate and Schmitt (2014) calls it an *ordinary* debt, in contrast with the *pathological* debt incurred to finance the purely monetary payment of country A's deficit. Two foreign debts, one ordinary and the other pathological, call for two foreign creditors; who are they? It is hardly conceivable to maintain that the credit accompanying the second debt is owned by the same residents of R who grant A the first, ordinary loan. It would also be difficult to pretend that the beneficiaries of country A's pathological debt are some other residents of R. Indeed, if this were the case economists would have been aware of it already, because the servicing of this second debt would have noticeably modified the

domestic financial situation of creditor countries. If the existence of the pathological duplication of indebted countries' external debts has so far gone unnoticed, it is because the second debt is incurred by country A not to R's residents but to a stateless financial bubble whose very presence has still to be satisfactorily explained by economists.

The nature of the sovereign debt crisis can be understood only once it is made clear that a country's indebtedness can only result from a process of duplication whereby the country running a deficit has to borrow an amount of foreign currency that it should obtain free of cost once net imports are paid by its residents. It is because money is transformed from a means into an object of exchange that countries get indebted. From the moment a monetary payment adds up to an equivalent real payment, duplication is unavoidable. This actually happens for the totality of countries' external payments, but to the extent that the double payment is reciprocal, it has no consequences on countries' external debts. If country A exports for ten MR to country R and imports from it also for ten MR, both country A and country R pay twice their reciprocal exchanges, but then a compensation intervenes, which cancels out their reciprocal debts: country A's pathological debt is compensated by country R's pathological debt. It is when country A's total imports (purchases) exceed its total exports (sales) that things go wrong, because today A has to borrow abroad a sum of foreign currency equal to the amount of its deficit. If the payment of country A in MR replaced the payment of its residents, no duplication would occur. Yet, this is not what happens: the two payments add up, the second, financed by a sum of MR borrowed by country A, adding to the first, carried out in MA by A's importers. As a result, country A incurs a totally unjustified debt, which is the exact definition of its sovereign debt.

The conclusion is that countries' sovereign debts are of a pathological nature and that a reform of the actual system of international payments is urgently needed to prevent the very formation of these debts. Countries' residents pay all of their foreign purchases and that should be enough. Any additional charge on countries is illegitimate and must be cancelled. This is what the reforms advocated in Chapter 11 are aiming at making sure.

10

Reforming Domestic Payment Systems

The global financial crisis that erupted in 2008 highlighted that banks do not make a clear distinction between the monetary and the financial intermediation they carry out, as explained in this volume. Indeed, in the early drafts of his *Treatise on Money*, Keynes (1973a: 91) noted that a bank is both a ‘money purveyor’ and a ‘credit purveyor’ within a monetary economy of production. This distinction has been lost in both economic theory and policy since then. It was, nevertheless, at the core of Ricardo’s (1824) *Plan for the Establishment of a National Bank*. As he observed, ‘[t]he Bank of England performs two operations of banking, which are quite distinct, and have no necessary connection with each other: it issues a paper currency as a substitute for a metallic one; and it advances money in the way of loan, to merchants and others’ (ibid.: 276). Since these two operations are conceptually distinct, Ricardo (1824: 276) explained that they can be carried out by two separate bodies, ‘without the slightest loss of advantage, either to the country, or to the merchants who receive accommodation from such loans’. In fact, as we will argue in this chapter, this functional separation is not only harmless but a structural factor of financial stability, because it allows to avoid that banks issue empty money in purely financial transactions that do not generate new income within the economic system as a whole. Let us expand on this, elaborating on Ricardo’s distinction between money and credit, to show that banks’ bookkeeping must also distinguish between income and capital in order to avert financial crises at systemic level. The next section will thus address the implementation of Ricardo’s distinction between money and credit in the ledgers of any banks (rather than at central bank level only). The second section will explain the importance of introducing a third department in banks’ accounting, beyond the two separate departments for monetary

and financial intermediation respectively. A fixed-capital department has indeed to account for those profits that are invested in the production of capital goods and which therefore should not be available any more in the form of bank deposits to finance lending operations on any kinds of markets (to avert capital over-accumulation and the resulting macroeconomic disorders that we explained in Chapter 7). The third section will present a stylized case to illustrate how the working of a bank's three-department bookkeeping makes sure that no systemic financial crises can happen again.

The separation of money and credit in banks' books

The separation of money and credit stems from a correct understanding of their nature. As we have pointed out throughout this book, neither post-Keynesian nor more orthodox economists managed to provide a completely logical understanding of money and credit in a monetary production economy. To be true, Fischer (1983: 4) recognized that '[b]anks do two things in this economy. First, they act as financial intermediaries. [...] Second, they provide transactions services, making payments as demanded by the households.' Nevertheless, this analytical distinction between financial and monetary intermediation that any bank carries out has not been translated in operational terms to date. Also, some endogenous-money theorists noticed that 'total spending is the sum of both incomes generated in "the circular flow" – which primarily finances consumption – plus the growth in debt – which primarily finances investment' (Keen 2011: 155). This quotation suggests that income moves in a circular flow: it is generated as a result of production and is destroyed as a result of consumption. The same quotation suggests also the existence of a link between (pathological) financial investment and banks' financial intermediation. In fact, the growth in debt (bank loans) observed in the run-up to the global financial crisis that burst in 2008 has been instrumental in driving up real and financial asset prices, and not so much to finance productive investment by non-financial firms.

[E]ntrepreneurs are not the only ones who borrow money: so do key actors in Minsky's explanation for Great Depressions, 'Ponzi Financiers'. These borrowers do not primarily invest with borrowed money, but buy existing assets, and hope to profit by selling those assets on a rising market. Unlike Schumpeter's entrepreneurs, whose debts today can be repaid from profits tomorrow, Ponzi Financiers

always have debt servicing costs that exceed the cash flows from the assets they purchase with borrowed money. They therefore must expand their debts or sell assets to continue functioning.

(*ibid.*: 156)

Keen (2011) is on the right track when he argues that bank credit may be used either to finance investment spending – hence generating a new income in the economic system – or to inflate an asset bubble that can increase the system's financial instability in the long run. However, he remains at the financial-behavioural level of investigation, rather than also considering the monetary-structural level, since he argues that '[t]he key early warning indicators [of a financial crisis] include the ratio of private debt to GDP, its rate of growth and acceleration, and sustained sectoral imbalances' (Keen 2013: 248). These indicators may reveal that a (systemic) financial crisis is building up and threatening the whole economic system. Nevertheless, they cannot explain the structural factors of such a crisis, unless they are considered within a monetary-structural framework as provided throughout this volume (see, in particular, Part III). Indeed, behaviour may impinge on financial stability, but cannot generate a systemic crisis, that is, a crisis that concerns the economic system as a whole, unless the system as such is affected by some pathologies, which affect its own structures and institutions rather than merely its different agents as generally maintained by the economics profession.

To make sure that the practical and conceptual confusion between money and credit will not cause financial instability and eventually a systemic crisis, financial regulators must adapt Ricardo's structural reform (see his *Plan for the Establishment of a National Bank*) to an endogenous-money banking system, where a number of non-bank financial institutions have further blurred the empirical evidence concerning money and banking. Rather than wasting time on a political debate over the best regulations to control agents' behaviour – which is in fact a never-ending task, because regulators will always run behind their preys – the framework that best serves the objective of systemic financial stability is to make sure that the emission of money will never be mixed up with the provision of loans, since the latter must be financed with a pre-existing income as a general rule. This rule should not be directed at agents' behaviour, however, because it would not be enough to repair a systemic failure, besides being easy to circumvent in a way or another – in particular, owing to 'financial innovation' and 'financial engineering' that a number of commentators in the footsteps of Warren Buffet have dubbed 'financial weapons of mass destruction'.

Therefore we propose to expand on Ricardo's two departments in banks' books considering the endogeneity of money, which Moore (1988: 372) characterizes as a magnitude that is 'credit-driven and demand-determined'.

Suppose that a given bank separates the recording of all its daily activities into an issue department (for money emissions) and a banking department (for financial intermediary transactions). The former department records any money emissions the bank carries out in the domestic payments system. Let us consider the payments that create new deposits in the whole banking system, and which derive from production. The payment of a wage bill will give rise to the entries recorded in Table 10.1.

Entry (1) shows the emission of money to the benefit of firm F, which has to pay out the wage bill (suppose for an amount of x money units, m.u.) to wage earners who are credited with a bank deposit by entry (1'). At the end of the day when this payment has been made, and if no other payment has occurred concerning either a new payment of wages or the expenditure of the income formed through payment (1), the bank transforms the monetary debt of the firm (entry 1) into a financial debt (entry 2'), on which interests will accrue daily as this is standard practice in any banks. The balance (*) of all these entries shows that the firm has indeed a financial debt to the bank, which in turn is financially indebted to wage earners. So far the result is no different from today's single-department banks' bookkeeping, which as a matter of fact already records both firms' financial debts and wage-earners' financial credits to the relevant bank.

Table 10.1 The result of the payment of wages through the two departments of a bank

Bank B			
Assets		Liabilities	
Issue department (I)			
(1) Credit on firm F	+x m.u.	Department II	+x m.u.
(2) Credit on firm F	-x m.u.	Department II	-x m.u.
(*)	0 m.u.		0 m.u.
Financial department (II)			
(1') Department I	+x m.u.	Wage earners' deposit	+x m.u.
(2') Credit on firm F	+x m.u.	Department I	+x m.u.
(*) Credit on firm F	x m.u.	Wage earners' deposit	x m.u.

Note: (*) is the balance of those entries that are recorded in the relevant department.

Now, when the bank carries out a purely financial-market transaction, one that does not generate a new income for the whole economic system, the two-department structure of bank accounts will show the maximum amount (of income) that the bank is entitled to spend – for its own sake or for its clients’ – without generating an inflationary increase in demand due to an excess of credit. As Table 10.1 shows, the bank can lend no more than x units of income on the financial market, as this is the total amount of purchasing power available in the whole economic system. The logical rule according to which the bank can neither lend nor spend on purely financial-market transactions more than what is deposited with it, will necessarily be enforced by distinguishing between the first two departments advocated by Schmitt’s 1984 reform. Indeed, the principle of double-entry bookkeeping applied to bank ‘departmentalization’ spares banks the risk of lending an amount greater than the one generated by production. The ‘golden rule’ applied today, according to which the sum lent must be backed by an equivalent bank deposit, is not enough to avoid this risk, because it does not impede the superposition of a monetary to a financial intermediation. If a mere sum of money takes the place of a positive income, the ‘golden rule’ may still be respected – the loan being balanced by a deposit of nominal money – but the excess of credit supplied by the bank is a source of inflation. By keeping the two departments separate, the reform of the domestic payment system avoids the financing of bank loans through money creation, the daily credit–debit relationship between the first and the second departments setting the limit to the loans the bank can supply during this period of time.

Let us suppose that some clients, C , of bank B obtain a loan of y units at some moment during the day. Like any other payment, the loan to C calls for the intervention of the bank both as a monetary and a financial intermediary. The sum lent to C reduces by y the debit–credit relationship between Department I and Department II, the balance, equal to $x - y$, determining the amount that the bank can still lend on the financial market. If no other payment intervenes during the day, the cancellation of the entries in the issue department and their transfer to the financial department will give rise to the situation as described by the end balance of Department II shown in Table 10.2.

Entries (1) and (1′) are the ones implied, respectively, in the emission of money and in the transfer of income involved in the loan granted by bank B to its clients. As a result of the loan, the debit–credit relationship between the two departments is reduced from x to $x - y$, as a result of entries (2) and (2′). The successive entries, (3), (4), (3′), and (4′) describe what happens at the end of the day, when all the entries still recorded in

Table 10.2 The result of a bank loan

Bank B			
Assets		Liabilities	
Issue department (I)			
Credit on firm F	x m.u.	Financial department (II)	x m.u.
(1) Financial department (II)	y m.u.	Clients C	y m.u.
(2) Credit on firm F	x m.u.	Financial department (II)	$x - y$ m.u.
(3) Financial department (II)	x m.u.	Clients C	y m.u.
(4) Clients C	y m.u.	Debit to firm F	x m.u.
(5)	0 m.u.	Financial department (II)	y m.u.
			0 m.u.
Financial department (II)			
Issue department (I)	x m.u.	Wage earners' deposit	x m.u.
(1') Wage earners' deposit	y m.u.	Issue department (I)	y m.u.
(2') Issue department (I)	$x - y$ m.u.	Wage earners' deposit	$x - y$ m.u.
(3') Credit on firm F	x m.u.	Issue department (I)	x m.u.
(4') Issue department (I)	y m.u.	Clients C	y m.u.
(5') Credit on firm F	x m.u.	Wage earners' deposit	$x - y$ m.u.
		Clients C	y m.u.

the issue department are transferred to the financial department. Finally, entry (5) is the end balance of Department I, while entry (5') is the end balance of Department II, which shows that, having been 'lent', as it were, part of the income saved by income earners, clients C are now enabled to purchase an equivalent part of the output stocked by F.

Let us now imagine that bank B is willing to grant a loan to some agent wishing to purchase an existing real-estate object. If this occurs, the bank will have to finance this mortgage loan with the (pre-existent) deposit that wage earners hold and will have to surrender in exchange for some financial assets, such as those assets that the bank provides through securitization of the underlying mortgage. We can also suppose that wage earners spend the rest of their deposit to purchase part of F's output. Table 10.3 shows the bookkeeping results of the relevant transactions in this case.

Entries (1) and (1') refer to the mortgage financed by bank B through the sale of financial assets to original deposit holders (wage earners) who dispose of their liquid balance to transform their wealth into a less liquid financial asset, possibly with a higher remuneration than their initial sight deposit, this deposit being transferred to the real-estate agent, through entries (2) and (2'), in payment of the real-estate object

Table 10.3 The result of a residential mortgage loan

Bank B			
Assets		Liabilities	
Issue department (I)			
Credit on firm F	x m.u.	Financial department (II)	x m.u.
(1) Financial department (II)	y m.u.	Residential mortgage	y m.u.
(2) Residential mortgage	y m.u.	Financial department (II)	y m.u.
(3) Financial department (II)	$x - y$ m.u.	Debit to firm F	$x - y$ m.u.
Credit on firm F	y m.u.	Financial department (II)	y m.u.
(4) Financial department (II)	y m.u.	Debit to firm F	y m.u.
	0 m.u.		0 m.u.
Financial department (II)			
Issue department (I)	x m.u.	Wage earners' deposit	x m.u.
(1') Wage earners' deposit	y m.u.	Issue department (I)	y m.u.
(2') Issue department (I)	y m.u.	Real-estate agents	y m.u.
(3') Wage earners' deposit	$x - y$ m.u.	Issue department (I)	$x - y$ m.u.
Issue department (I)	y m.u.	Real-estate agents	y m.u.
(4') Credit to firm F	y m.u.	Issue department (I)	y m.u.
(6) Credit to firm F	y m.u.	Real-estate agents	y m.u.

sold to the mortgage holder. Entries (3) and (3') concern the expenditure by income holders of the rest of their deposit ($x - y$) for the final purchase of part of F's output. At the end of the day, the balance of the issue department is transferred to the financial department, as shown by entries (4) and (4'), and entry (5) shows that the real-estate agent has replaced wage earners as owner of the income required for the final purchase of the output still stocked with firm F.

As these examples make clear, the 'financial circulation' of bank deposits preserves the purchasing power of income produced on the labour market, rather than eroding it through a series of bubbles that have up to now been generated by banks moving 'forward in step' (Keynes 1930/1971: 23) – which occurs when each of them increases its *ex-nihilo* loans at the same rate of growth as any other bank, thereby making it impossible to detect any financial problem through interbank settlements during or at the end of the business day.

The separation of income and capital in banks' books

The separation of money and credit is a necessary structural reform in banking, in order to avoid the occurrence of systemic crises. This is not

sufficient, however, to make sure that such crises will not occur. Once the structure of banks' bookkeeping is reformed to reflect the essential distinction between money and income, it is necessary to elaborate on it also to distinguish income from capital. This is so because of the macroeconomic nature of capital goods: being the result of investment, capital goods encapsulate agents' savings, whose income is thus 'fixed' within these goods as a result of their production (see Schmitt 1984a). In other words, when firms dispose of their profits to finance a new investment, the corresponding income is transformed into fixed capital. It would thus be problematic, as is currently the case, if the bank deposits corresponding to firms' invested profits remain in the financial department of banks' bookkeeping: in that case, indeed, the bank where these deposits are recorded would be in a position to lend (as it does to date) the corresponding amount to any kind of borrowers, although the object of these deposits is an income that has been transformed into fixed capital for ever and that, as a result, should not be anymore available to finance any kind of transactions. In fact, banks lend this macroeconomic saving (an income that society as a whole has invested in the production of capital goods), and earn an interest from borrowers, which more than compensates the interest paid to the corresponding deposit holders.

As clearly stated by Schmitt in his 1998–99 manuscript on unemployment, the present pathological regime of capitalism is characterized by the loan of the deposits (formed by the expenditure of profits) granted by banks to households. It is because of this loan that households lose the ownership of fixed capital, which is appropriated by the set of disembodied firms. Unlike what happens to the income spent for the final purchase of wage-goods, the expenditure of profit generates a deposit that will never be destroyed. 'To the extent that the macroeconomic costs of production are overcome, *deposits* obey to the law of "circulation": they circulate indefinitely in the national economy' (Schmitt 1998–99a: 98, our translation). It is the loan of these deposits that defines capitalism. 'What characterizes most profoundly the capitalistic mode of production is the fact that the deposits, formed at the moment profits of any kind are created, are lent by B, set of banks (or banking system)' (ibid.: 100, our translation). The loan of those bank deposits that correspond to profits is pathological because, since they derive from the expenditure of wages, profits should never finance any other expenditure. If the deposit of profits feeds a new loan, the income already spent by wage earners is spent again by borrowers, which means that the loan of the deposit of profits allows an income to be spent twice. Finally, in today's

pathological system profits are themselves pathological and add up to wages, as is the case for inflation profit in Keynes's *Treatise on Money*.

Schmitt's reform radically changes this state of affairs and will make sure that profits, of any kind, will stem from wages and will therefore not be issued on the financial market through a loan granted by banks to households. His structural reform of banking aims at securing that all invested profit (that is, an income transformed into fixed capital) is booked in a separate department of banks: a fixed-capital department has then to be introduced in banks' bookkeeping. The role of the fixed-capital department is to remove the deposit of profits from the financial department and so avoid that those deposits become the object of banks' financial intermediation.

During the period of time separating the formation of a profit from its final [investment], the corresponding deposit is lent by the banking system: it is this loan of an income already extinct, because it is derived from *spent* wages, which is deleterious. The deposit of profits must remain pending, in its original state, without the possibility to lend it, until their holders will decide to [invest] it.

(*ibid.*: 135, our translation)

The transfer of the deposit of profit from the second to the third department will make it impossible for profits to be spent within the payment of wages, which is what explains why, in the current pathological system, fixed-capital goods are appropriated by disembodied firms and wage earners obtain nothing in exchange for their labour. Indeed, capitalism 'is the regime where firms pay part of labour through the expenditure of a positive purchasing power, previously "seized" or "captured" from households' (Schmitt 1998–99b: 24, our translation). The transfer to the fixed-capital department of the deposit of profit has the effect of avoiding the inclusion of the expenditure of profits in the payment of wages and, by the same token, the loan of the deposit of profits to households.

Let us elaborate on this considering the initial example. Suppose, in this regard, that the actual deposit holder (a real-estate agent) recorded at the end of Table 10.1 has disposed of this deposit (x m.u.) on the market for produced goods and services. As a result, firm F obtained the ownership of the corresponding amount when it sold, with a markup, part of its output of consumption goods at a price that more than covers the relevant costs of production. Assuming that the ensuing profit amounts to say y units of money, this is the amount that

Table 10.4 The entry of profit in the second department

Bank B			
Assets		Liabilities	
Financial department (II)			
(1) Stock of goods	y m.u.	Debit to firm F	y m.u.

firm F supposedly will invest for the production of fixed-capital goods in a subsequent period. Table 10.4 shows the balance of the financial department at the end of the day.

Entry (1) shows that firm F is the holder of a bank deposit equal to y , its profit, and that the object of this deposit is the stock of goods financially owned by the bank. If F's profit were to remain available in the financial department, F's subsequent investment would entail its expenditure and the formation of a new deposit that would continue to exist in time, forever. In order to avoid the formation of such a perennial deposit that would feed a constantly renewed loan to households, Schmitt's reform envisages the transfer to the third department of the income formed to the benefit of F. At the end of the day, entries in the financial and in the fixed-capital departments would be as shown in Table 10.5.

Entries (1) and (1') relate to the transfer from Department II to Department III of F's profit, while entry (2) is the end balance of the financial department, where F's profit is no longer available for lending on the financial market. If, as supposed, firm F is going to invest its entire profit in the production of investment or fixed-capital goods, the payment of wages will take place according to the same procedure described in

Table 10.5 The transfer of profit to the third department

Bank B			
Assets		Liabilities	
Financial department (II)			
Stock of goods	y m.u.	Debit to firm F	y m.u.
(1) Credit on firm F	y m.u.	Fixed-capital department (III)	y m.u.
(2) Stock of goods	y m.u.	Fixed-capital department	y m.u.
Fixed-capital department (III)			
(1') Financial department (II)	y m.u.	Debit to firm F	y m.u.

Table 10.6 The investment of profit

Bank B			
Assets		Liabilities	
Issue department (I)			
(1) Credit on firm F	y m.u.	Financial department (II)	y m.u.
(1) Financial department (II)	y m.u.	Debit to firm F	y m.u.
(2) Credit on wage earners	y m.u.	Financial department (II)	y m.u.
(2) Financial department (II)	y m.u.	Debit to wage earners	y m.u.
	0 m.u.		0 m.u.
Financial department (II)			
Stock of consumption goods	y m.u.	Fixed-capital department (III)	y m.u.
(1') Credit on firm F (stock of fixed-capital goods)	y m.u.	Wage earners' deposit	y m.u.
(2') Wage earners' deposit	y m.u.	Stock of consumption goods	y m.u.
(3) Stock of fixed-capital goods	y m.u.	Fixed-capital department (III)	y m.u.
Fixed-capital department (III)			
(3') Financial department (II)	y m.u.	Debit to firm F	y m.u.

Table 10.1 and, instead of implying the expenditure of F's profit on the labour market, will produce a new income with fixed-capital goods as its real content. Generalized exchange will then enable wage earners to purchase the stock of consumption goods still unsold, while the sum invested by F is transformed into a macroeconomic saving deposited with Department III (Table 10.6).

The payment of wages for the production of fixed-capital goods is recorded as entries (1) and (1') in the issue and in the financial departments, whereas entries (2) and (2') describe what happens when wage earners spend their income for the final purchase of the stock of consumption goods produced in a previous period. The end result, that is, entries (3) and (3'), shows that the stock of fixed-capital goods obtained through the investment of profit is still owned by households (in particular by F's shareholders). If F's profit had not been transferred to Department III, it would have been spent on the labour market, and its expenditure would have defined the appropriation of fixed-capital goods by the depersonalized firm (see Chapter 7). Schmitt's reform avoids this pathological appropriation by safeguarding the financial link between households' macroeconomic saving and the fixed-capital goods produced by F, as described by entries (3) and (3') in Table 10.6.

The rule according to which any profit entered in the bank's financial department is automatically transferred to the fixed-capital department at the end of the day does not admit of any exception and does not depend on the amount of profit firms are willing to invest. In fact, firms are in no way forced to invest the totality of their profits. In our example, if firm F were to change its plan and decide to redistribute part of its profit, say z , a sum of z would be transferred back from the third to the second department and distributed as interest and/or dividend to creditors and/or shareholders, thus reducing the amount deposited as fixed capital into Department III. Finally, only the amount of the profit invested in the production of fixed-capital goods, $y - z$, will remain deposited in the fixed-capital department and define the economy's macroeconomic saving (Table 10.7).

Entries (1) and (1') concern the end situation of day one, when F's profit is transferred to Department III and the production of fixed-capital goods is still undetermined. The stock of goods available at the end of the first day is thus simply a stock of profit goods. It is only on day two that it will be possible to determine whether part of this stock is made up of interest and dividend goods or not. In our example, F's decision to redistribute z units of its profit reduces to $y - z$ the amount of fixed-capital goods corresponding to the fixed capital deposited with Department III. F's creditors and shareholders become the holders of an income of z units deposited with the financial department, as shown

Table 10.7 The entry of redistributed and invested profit

Bank B			
Assets		Liabilities	
Financial department (II)			
(1) Stock of profit goods	y m.u.	Fixed-capital department (III)	y m.u.
(2) Fixed-capital department (III)	z m.u.	F's creditors and shareholders	z m.u.
(3) Interest and dividend goods	z m.u.	F's creditors and shareholders	z m.u.
(4) Stock of fixed-capital goods	$y - z$ m.u.	Fixed-capital department (III)	$y - z$ m.u.
Fixed-capital department (III)			
(1') Financial department (II)	y m.u.	Debit to firm F	y m.u.
(2') Credit on firm F	z m.u.	Financial department (II)	z m.u.
(4') Financial department (II)	$y - z$ m.u.	Debit to firm F	$y - z$ m.u.

by entries (2) and (2'), and are thereby enabled to purchase an equivalent part of the initial stock of profit goods, as described by entry (3). Once interest and dividend goods are purchased, only entries (4) and (4') remain, the financial relationship between the second and the third department defining the amount of fixed capital formed in the whole economic system.

Being transferred from the financial into the fixed-capital department of banks' books, the deposits corresponding to the investment of profits cannot give rise to capital over-accumulation when firms' fixed capital is amortized (see Chapter 7). The amortization of capital goods implies a production of replacement goods, whose producers, wage earners, are paid an amount of (say z) money units that, previous to the structural reform suggested in this chapter, are devoid of any original purchasing power, because investment goods are owned, in fact, by disembodied firms as explained in Chapter 7. Once the distinction between the financial and the fixed-capital departments in banks' bookkeeping has been introduced all newly formed investment goods will be owned by the firms' shareholders rather than by pathological capital. As a result, fixed-capital amortization will give rise neither to inflation nor to unemployment, contrary to the current monetary disorder that originates in the lack of distinction between money, income, and capital (see Chapter 7). Let us consider a stylized case in the remainder of this chapter to illustrate the merits of such a monetary-structural reform of banks' bookkeeping as regards the avoidance of an utterly damaging systemic crisis of 'globalized' finance.

A structural solution to avert systemic financial crises

Financial setbacks will continue to occur, as a result of different behavioural factors, such as mismanagement, wrong risk allocation, and agents' interconnectedness. This is in the nature of human beings and human-based institutions, and is indeed unavoidable – even though financial regulators as well as financial education can do much to limit both the number of financial setbacks and the negative impacts of them on the population at large. Be that as it may, what must absolutely be averted is another *systemic* financial crisis like the global financial crisis that burst in 2008 after the collapse of the housing market in the United States and the related problems in the mortgage-loan and securitization industries, which ramified across the global economy.

Generally speaking, economists identify in the activity of speculators and unscrupulous financial operators, in a framework characterized

by insufficient regulatory norms, the cause of financial crises. What is missing in their analysis is the fact that speculation, in all its various forms, exists because of the existence of a speculative financial capital of a pathological nature. Even though they recognize the existence of what has been called a financial bubble, they fail to explain how this bubble can form in the first place. Because they lack a clear-cut conceptual distinction between money, income, and capital, mainstream economists are unable to grasp the pathological process leading to the creation of a purely nominal capital deprived of any real content. Now, it is because of the presence of such a capital that financial speculation has never stopped increasing in the last decades. Financial speculation would not even exist or would be insignificant if no pathological capital were available. The occurrence of financial crises, of inflation and unemployment, is ultimately best explained by the pathological process of capital accumulation that characterizes the present capitalist regime. This is indeed so because that pathological process – which is essentially due to a lack of conformity of the present system of national payments with the macroeconomic laws deriving from the logical relationship between money, income, and capital established by production – is at the origin of a ‘duplication’ that feeds inflation and is at the same time the source of deflation (see Chapter 7).

Once Schmitt’s monetary–structural reform advocated in this chapter is implemented within domestic payment systems, it will not be possible any more to blur the distinction that, essentially, exists between money and credit, which fuels inflation and eventually unemployment through capital over-accumulation. The three essentially distinct departments introduced by the reform in banks’ bookkeeping are instrumental in order to reach this goal. Systemic financial stability will thus be the natural outcome of a payment architecture that works in line with the nature of money, income, and capital. Let us illustrate this in this section, considering a stylized case implying two banks within the same (domestic) payment system reformed along the lines explained above.

It is well known that banks are always keen to open reciprocal credit lines in the interbank market, whose amounts are generally titanic, short-term, and uncollateralized – except when a financial crisis exists or is looming. Let us therefore assume that two banks, A and B (the latter representing the whole banking sector apart from the former), agree to set up a reciprocal credit line of x units of their domestic currency, considering that in case of a (systemic) problem the central bank will intervene to avoid a financial crisis. To date, there is no endogenous limit for any bank in opening such a credit line, owing

to money's endogeneity and the blurring of the distinction between money and credit in banks' bookkeeping as explained above. In a structurally reformed payment system, by contrast, each bank will be credit constrained in any non-income generating transaction, by the amount of income (in the form of bank deposits) that exists in the whole banking system at that time.

Logically, in a monetary economy of production, income must first be produced (on the labour market) in order for banks to provide a credit line (to any kind of agents) for non-income generating transactions (on either goods or financial markets). Our stylized case should therefore consider first the payment of the wage bill for the relevant production – a transaction that we already explained with Table 10.1. Hence, if bank B wants to open a credit line to bank A, the maximum amount that the latter may borrow is given by the amount of income recorded as a bank deposit of wage earners (x m.u.). However, bank A is not in a position to grant a credit to bank B, unless it also records some deposits as a result of the compensation for any kinds of production activities through the payment services it provided to the relevant (financial or non-financial) firms. This is already an important monetary-structural constraint on banks' credit provision for non-productive transactions.

Further, once all bank deposits corresponding to invested profits are moved into a fixed-capital department of banks' bookkeeping, a second important monetary-structural limit on banks' credit provision for non-productive transactions will operate. No bank will be in a position to lend the deposits corresponding to fixed-capital accumulation. Hence no bank will originate a loan for non-income-generating expenditures that does not rely on a pre-existent purchasing power in the form of a deposit within the (domestic) payment system.

On the whole, only production will allow banks to provide credit lines that do not rely on pre-existent bank deposits. The emission of money will always be credit-driven and demand-determined (see above), but there will be an explicit recognition of the amount of newly issued money (which increases the sum total of deposits in the whole banking system) as well as an updated record of fixed-capital accumulation in banks' books.

11

Reforming the International Monetary System

As we have seen in Chapter 9 and as widely recognized by economists all over the world, '[t]he international monetary and financial systems are clearly in trouble, and reforms are called for' (Dooley et al. 2009: 4). This call for reforms of the international monetary regime is far from new. Before and after the Bretton Woods conference, where the discussion centred on Keynes's and White's proposals, several plans of reform were proposed and debated, among these Schumacher's (1943a, 1943b) and Stamp's (1963). With the partial exception of White's plan, which *de facto* merely enabled the passage from the sterling exchange standard to the US dollar exchange standard, none of these plans was retained and the world went on using what was to be universally labelled a non-system of international payments. The need for a reform capable of eradicating the international causes of monetary instability and financial crises has grown stronger year after year and has now reached a critical urgency. However, as Dooley et al. (2009: 4) emphasized, 'in weighing potential dimensions of reform, there remains considerable uncertainty and debate about the relative importance of factors that have driven and continue to drive the current crisis'. Indeed, unless the origin of international monetary disorder is correctly individuated, no successful plan of reform can be envisaged. Now, quantum monetary macroeconomics provides a new analysis of monetary and financial disorders, and is equal to the task of suggesting a solution that can take three alternative forms: a reform concerning any single country, a reform concerning a group or a community of countries, such as the EU, and a reform of the entire system of international payments.

In the first section of this chapter, we will present the reform recently elaborated by Schmitt (2014) on the basis of his analysis of the pathological formation of countries' sovereign debts that we have summarized

in Chapter 9. The interest of this 'individual' reform is that it can be implemented by any single country irrespective of what is done by the rest of the world and without causing any disruption. Thanks to this reform, the country implementing it would be able to avoid the present pathological duplication of its external debt, and its government's budget would earn the domestic income that is lost today, because of the net expenditures carried out by its residents. By reducing the total charge of the deficit run by the country adopting it to a one time only charge, this reform would allow foreign creditors to be paid their due and, at the same time, it would enable the debtor country to maintain its level of employment unaltered, which today is decreased by the second charge of the country's net foreign purchases.

This reform is not immediately and easily understood. However, its advantages are so important and its implementation so easy that it deserves our attention. Indebted countries all over the world would greatly benefit from the possibility to avoid the monetary and financial disorders deriving from the payment of their transnational transactions, without any need to wait for an international conference to ratify a worldwide agreement on a universal system reform.

In the second and third sections of this chapter, we will investigate the principles and steps that must be followed to elaborate and implement a reform either at a 'multinational' or at a world level.

Among the proposals most seriously discussed in recent years by experts in the field of international payments we find those based on the use of Special Drawing Rights (SDRs) and those advocating the creation of a new international payments system along the lines of the plan presented by Keynes at the Bretton Woods conference (1944). Both these proposals are unsatisfactory since they cannot be reconciled with the flow or vehicular nature of money and since they advocate the use of an international reserve asset backed by national currencies or by assets of national central banks. In the case of SDR-based reforms, the *Report of the UN Commission of Experts* suggests the use of a 'new global reserve currency that could be managed by the IMF or by a new institution – a Global Reserve Bank' (D'Arista and Ertürk 2010: 74) either in a 'world-wide system of swaps among central banks with the contributions in their currencies as backing for the global currency' (ibid.: 75) or through 'the commitment of member countries to accept it [the new reserve currency] in exchange for their own currencies' (ibid.: 75), or by authorizing the Global Reserve Bank 'to use them to buy government securities or lend them, providing backing for the global currency in the same way national currencies are backed by the assets of national central banks'

(ibid.: 75). In every one of these three alternatives, SDRs are conceived of as positive assets deriving their value from national currencies or from the assets deposited with national central banks. The emission of SDRs is identified with that of an asset and the Global Reserve Bank is seen as an institution that 'must have the authority to create credit and must use some form of backing that can channel credit to the recipients' (ibid.: 75). This is to forget, however, that banks cannot create credit, neither at the national level nor at the international level. The idea of credit creation is totally irrational and theoretically as well as empirically unfounded.

The proposal for the creation of a new international agency working along the lines of Keynes's International Clearing Union (ICU) does not avoid the problem, since it is supposed to 'clear cross-border transactions in members' own currencies by crediting and debiting their clearing accounts' (ibid.: 77) and since it is assumed that '[t]hese clearing accounts would, in fact, constitute the international reserve of the system, held by the ICA [International Clearing Union] and valued using a trade-weighted basket of members' currencies' (ibid.: 77). Even Keynes's original plan is not satisfactory, both because the new international currency to be issued by the ICU, the *bancor*, is defined in terms of national currencies and gold, and because payments in *bancor* can never be final. What Keynes's plan fails to recognize is that money is a *means* of payments and not their *object*. Nonetheless, some of Keynes's insights can be very fruitful if inserted in a theoretical framework where money, both national and international, plays the role of a numerical intermediary allowing for the final payment of domestic and external transactions.

Besides showing that any single country can protect itself against the monetary and financial disorders caused by the present non-system of international payments, the aim of this chapter is to show that the passage to an orderly system of international payments is possible. In particular, the European Union (EU) could implement easily enough a reform preventing its member countries to suffer from the serious drawbacks caused by the actual lack of finality of their external payments and from their sovereign debt crisis. An institution like the European Central Bank (ECB) can almost effortlessly be modified operationally in order to act as a true monetary and financial intermediary for any transaction carried out by EU's residents outside their own country. Such a 'regional' solution could also be adopted by other groups of countries and, eventually, generalized at a planetary level. What is certain is that a solution is needed, today more than ever, and that the future of our

economies depends on the reform that will be adopted and on how long we will have to wait for it.

A single country reform

The reform suggested here is the one conceived and developed by Schmitt in his 2014 paper on the sovereign debt crisis and its solution. Our presentation is derived from his, and attempts at summarizing it without simplifying it too much.

Let us go back to the case analysed in Chapter 9, where we considered a representative country, A, running an overall deficit being confronted with the rest of the world, R. Today, A covers its deficit, that is, the difference between its total expenditures or (commercial and financial) imports and its total receipts or (commercial and financial) exports, by borrowing abroad an equivalent amount of foreign currency. The foreign or external debt thus incurred by A should be equal to the amount borrowed to finance its deficit. Yet, a pathological mechanism due to the absence of a consistent system of international payments duplicates A's external debt: country A has to borrow twice, once to recover – through a foreign loan – the domestic income spent by its residents, and once to pay its creditors (residents of R) in foreign currency. The purpose of the reform is to avoid this duplication, that is, to provide A, for free, the foreign currency it needs to convey the real payment of its deficit from its residents to R's residents.

In order to facilitate the understanding of that fundamental principle of the reform, which is to guarantee the respect of the balance-of-payments identity between each country's global exports and imports ($IM \equiv EX$), let us refer to another proof of the double charge of countries' external debts proposed by Schmitt in his 2014 paper. Once again we consider country A and country R and we suppose that A's total expenditures, equal to 14 dollars, exceed its total receipts by four dollars and that A covers its deficit by borrowing four dollars from R. Since A's imports are R's exports and vice versa, we are tempted to conclude that A's imports are greater than R's imports, $IM_A > IM_R$, and that their equality can only be established over two or more periods on condition that country A succeeds in realizing, in the future, an excess of exports over its imports equal to its present net imports. If this were indeed the case, we would have to conclude that:

1. The balance-of-payments identity is no identity at all.
2. Country A does not pay for its net imports in the period when they occur.

In reality, things happen in an entirely different way, consistent with respecting the balance-of-payments identity. A's exports, initially equal to 10 dollars, are brought to the level of the country's imports, equal to 14 dollars, as a consequence of the payment of net imports carried out by its residents. By transferring to R the ownership over an equivalent domestic income, A's residents pay their net purchases of R's real goods through a sale of part of their own domestic real goods.

It is through the loan of four dollars granted by R's residents that the equilibrium between A's and R's real imports is established. By lending four dollars to A's economy, R's residents purchase an equivalent real production of A. It is true that foreign lenders, residents of R, will obtain only later, in a subsequent period p^* , the real products of R, because it is only in p^* that they will be reimbursed. Yet, it is also certain that it is since the initial period, p , that they acquire a right over A's real production. Indeed, it is since period p that R pays for the goods that it will obtain only in p^* . It is therefore since p that it brings its imports to the level of A's (14 dollars). By the same token, A equalizes its exports with its imports by increasing the former through the sale, taking place in period p , of the real goods that it will produce in p^* .

A's first payment of its net imports occurs in period p and consists in the transfer of an equivalent amount of A's domestic income. However, a second payment adds to the first, because R's exporters are still to be paid for their net foreign sales. When it initially borrows four dollars, country A gives to its foreign lenders the ownership over an equivalent amount of its domestic production. What happens now is that the four dollars thus borrowed are spent to pay R's exporters, which leaves A with the necessity to find four dollars more to reimburse the initial loan. The payment of R's lenders and the payment of R's exporters are two distinct payments that add to one another: the total charge of A's net imports is twice as high as it should be.

Since the formation of A's external debt, that is, from the moment A borrows four dollars abroad, the economy of A gives up a domestic income worth four dollars. This is so because the debt incurred by A defines precisely the fact that A owes R part of its domestic product. If the four dollars obtained by A were still at its disposal, nothing would be wrong, A's external debt being matched by the credit, worth four dollars, on R's own current output. This, however, is not what happens. The dollars borrowed abroad by A are necessarily spent to pay for A's net purchases or imports, and the country is thus deprived, without any real counterpart, of a domestic output worth four dollars.

If we have proposed once more Schmitt's (2014) analysis of the pathology affecting the payment of a country's net imports, it is because it is essential to the understanding of his reform. What we would like to emphasize at this stage is that *it is the payment of the deficit through a foreign loan that engenders the need for a second payment and thus duplicates the country's external debt*. The two payments are strictly correlated and so are the two debts financing them. Once the first occurs, the second is unavoidable. The duplication settles in from the moment A has to borrow abroad a sum of foreign currency despite the fact that its residents have already fully paid their net imports. Hence, any reform that attempted to avoid the duplication without neutralizing A's net foreign borrowing would be doomed to failure. One first, fundamental principle the reform must comply with is therefore that of avoiding the very formation of country A's external debt. This might seem difficult to accept, because it is apparently straightforward to infer that a country running a deficit has necessarily to incur a debt in order to finance its net imports. Appearances may be delusive, however, and in this particular case they certainly are, because it is illogical to pretend that country A has also to pay, additionally and through a foreign loan, what has already been paid by its importers. What Schmitt's (2014) analysis tells us is that external or sovereign debts, that is, debts incurred by the country considered as the set rather than the sum of all its residents, should not exist. They are pathological results and as such they should never be allowed to form in the first place.

The second principle is closely related to the first and states that the net importing country must not lose even the smallest part of its domestic income. What leads today to the formation of a country's external or sovereign debt is that this country is forced to borrow abroad in order both to recover the ownership over its domestic income and to pay its creditors in a foreign currency. The reform must sever the link between the payment in domestic income and the payment in foreign currency. In our example, the payment carried out by country A's residents, in MA, must be kept separate from the payment to the benefit of country R's exporters, in MR. In other words, international payments have to be internalized, so that net exports are paid through the expenditure of an income generated in the exporting country itself on behalf of the importing country, which, in exchange for the real goods imported in excess has to pay an equivalent part of its own real exports. The domestic income spent by A's residents must remain available within country A. In Schmitt's (2014) reform this result is achieved through the implementation of a *sovereign Bureau* charged to

1. pay A's exporters and be paid by A's importers in MA; and
2. pay R and be paid by R in MR.

Let us analyse these two functions of A's Bureau in succession.

Within the equilibrium between A's overall sales (exports) and purchases (imports), the Bureau simply pays A's exporters with the income spent by A's importers. As for A's net purchases, the domestic income spent in excess by A's importers is obtained by their Bureau as a net gain. No one except A's sovereign Bureau will be entitled to the income spent in excess by country A's residents.

Now, if the Bureau were to act as a mere intermediary between A's residents and R, that is, if it were charged to convert their payment in MA into a payment in MR to the benefit of R, no net gain would ever form for its benefit. Indeed, the Bureau's intermediation would not be enough to avoid A's double charge of its net foreign purchases. If the Bureau were to borrow abroad a sum of foreign currency in order to pay for the net foreign purchases of A's residents, the same situation as today would occur, two loans being necessary, one to pay R and one to avoid the loss of a domestic income. What has to be clearly understood is that the ownership over part of A's domestic income is lost from the very moment a foreign loan finances the payment of A's net foreign purchases or imports. The intervention of the Bureau as an intermediary does not substantially change what actually happens in the existing non-system of international payments. If the Bureau is to obtain as a *pure gain* the sum of income spent in excess by A's residents, the payment of R must not be carried out through a net foreign loan, and no direct link must exist between the payment in MA and the payment in MR.

The emergence of a net gain for the Bureau is an all-important result that can be attained only if the reform succeeds in avoiding the very formation of the country's external debt while guaranteeing the full payment of its foreign partners. The second function the Bureau has to fulfil is therefore essential to the fulfilment of the first. R has to be paid in MR without A incurring a net external debt. How can this be done, given that A's receipts in MR are lower than its expenditures? As the reform can be implemented by any single country, the answer cannot rely on the intervention of a supranational bank, which implies that A cannot avoid borrowing abroad a sum of MR. Thus, the problem has apparently no satisfactory solution: having to borrow from R the money it does not obtain through its foreign sales (exports), country A gets necessarily indebted to R. Well aware of this difficulty, Schmitt (2014) overcomes it by demanding that the Bureau, which borrows abroad on behalf of

country A, lend to R an amount of money R equivalent to the amount it borrows from R. The debt incurred by A is thus immediately matched by an equivalent credit, leaving unaltered the level of A's external debt, which, once the reform is implemented, will no longer increase at all. The immediate lending of A to R neutralizes therefore the very formation of A's external or sovereign debt while allowing its Bureau to obtain, as a net gain, the domestic income spent by its residents to pay for their net imports.

What remains to be explained is how R is paid for its net exports. This can be done easily by observing that the loan granted by A's Bureau to R has exactly the opposite meaning from the loan granted by R to A's Bureau. When A borrows from R, it is R that pays for part of A's net imports of R's current production: by lending to R the same amount A borrows from it, A pays for an equivalent part of R's imports. Finally, R obtains part of A's current production in exchange for its net exports. Both A and R pay in real terms their foreign transactions, the real goods purchased by R on behalf of A being perfectly matched by an equivalent purchase of real goods carried out by A on behalf of R. Let us repeat this crucial point: by lending to R, it is A itself that pays for an equivalent part of its exports, because it is A that gives R what R needs to pay A. The final payment defines an exchange between A's and R's real goods that takes place without forcing A to incur a net debt. This is to say that Schmitt's reform enables A to carry out its real payments without having to purchase the vehicular money required to convey them to R.

A difficulty remains: the loan obtained from R being used to finance an equivalent loan to R, a second foreign loan is necessary for A to finance its net foreign purchases. It thus seems that nothing has really changed and that A's deficit is again financed through an external loan that unavoidably increases its external debt and brings about the pathological duplication the reform was supposed to avoid. Even though the loan granted by A to R compensates the initial loan granted by R to A, it would be wrong to claim that they simply cancel each other out as if nothing had happened. This is not so, because the residents of R who grant the first loan to A are not the same as those to whom the loan is granted by A's sovereign Bureau. The fact remains that the second loan A has to incur increases its external debt apparently. In fact, as shown by Schmitt (2014), this is not the case, because the loan granted by A to R has the double effect to cancel simultaneously A's first and second debt. The decisive argument is not easy to grasp, and consists in saying that the two loans obtained by A, of 4 MR each in our numerical example, do not add to one another, because the second is nothing other than the

repetition of the first. The two loans are identical, because 'the second loan generates for a second time the first' (Schmitt 2014: 18, our translation). The credit obtained by A owing to its loan of 4 MR to R is thus sufficient to cancel the unique debt, also equal to 4 MR, corresponding to the two loans obtained by A: 'the loan granted by country A compensates both the second and the first loan to country A' (ibid.: 18, our translation). In other words, the loan that A grants to R cancels the very formation of the first debt incurred by A and not simply its result. The second loan merely recreates the initial debt so that the two loans of four MR lead to a total debt of four MR, which is entirely compensated by A's credit toward R. Finally, A obtains the units of MR (4 MR) required to convey to R the real payment of its net imports without incurring any net debt.

The functions of the Bureau

Let us present here, synthetically, the main functions of the sovereign Bureau.

1. The Bureau of country A, which is essentially a resident of A, will be paid by A's importers and will pay A's exporters, in MA, for all their foreign transactions. All the payments concerning A's importers and exporters will therefore take place between residents.
2. The Bureau will be charged to pay A's foreign partner countries (R) in MR and it will be credited by the payments of R, also in MR.
3. The Bureau will lend to R a sum of MR equal to that of country A's overall imports. This is a key function with the goal to avoid the financing of A's net expenditures by a foreign loan, and will take place automatically any time, in a given period, country A's total foreign expenditures exceed its total foreign receipts.
4. The Bureau will obtain as a net gain the difference between the amount of income A spent by A's importers and that paid to A's exporters, that is, the sum of domestic income corresponding to country A's net foreign purchases.
5. The fifth function concerns the use of the Bureau's net gain in MA. According to Schmitt (2014), this amount of income A has to be invested, either by the Bureau or by A's government, in a new domestic production in order to avoid a decrease in employment due to the fact that part of A's current domestic production is owned by R (following the loan granted by A's Bureau to R). One cannot overstate all the advantages that such a use of the Bureau's net gain would entail for the countries that today suffer the most because of the loss

of domestic resources caused by the double charge of their external debts.

The operational functioning of the sovereign Bureau

Let us conclude the presentation of Schmitt's 2014 reform by summarizing his practical example of the way the Bureau of country A would act following a repeated deficit of its overall foreign transactions. The numerical example we will refer to is the one we have already considered before: A imports for 14 units of MR and its exports are equal to 10 MR, and the exchange rate between MR and MA is $1 \text{ MR} = 1 \text{ MA}$. The reference time frame or period is the month and we suppose that A's deficit is reproduced month after month. In the first period, January, the Bureau obtains 14 MA from A's importers and pays 10 MA to its exporters. The Bureau's net gain is equal to four MA and matches the foreign loan it has to incur in order to finance the difference between A's imports and exports. At the end of January, A's external debt is positive and equal to four MR.

In February the Bureau borrows again four MR abroad but it also lends an equal amount, four MR, to R. The loan granted to R balances the loan granted by R in January and cancels out A's initial external debt. In the meantime A's Bureau gains, as a net profit, the sum of income A paid by A's importers in excess to the sum paid to A's exporters: $14 \text{ MA} - 10 \text{ MA} = 4 \text{ MA}$. Now, A's sovereign Bureau still needs to borrow four MR to cover the difference between country A's sales (10 MR) and purchases (14 MR). This second foreign loan is indeed necessary, because the first sum borrowed by A is entirely neutralized by the loan granted to R by A's Bureau. Yet, it would be wrong to infer that A's external debt thus increases by four MR with respect to the debt incurred in January and reaches the total level of eight MR. Indeed, both the first and the second loan (which is the mere repetition of the first) obtained by A are compensated by the loan granted to R so that, at the end of February, A's external debt is still equal to four MR only.

In March as well as in the following months, if A's overall deficit is still of four MR in each month, A's sovereign Bureau keeps earning four units of domestic income (expressed in MA) as a net gain, which it keeps investing in order to avoid a reduction in employment. At the same time, the Bureau continues to grant a loan of four MR to country R, which reduces A's external debt to zero, the debt of four MR formed in February being cancelled by the credit obtained in March. Finally, by subscribing a second loan of four MR A's sovereign Bureau re-establishes the level of A's external debt at four MR. This means that, as long as

A were to run a deficit equal to four MR *in each period*, month after month its external foreign debt would remain equal to four MR. Asymptotically, A's external debt would therefore be zero, in compliance with what the reform called for by the pathological nature of sovereign debts requires.

The 'multinational' and the world reforms

Unlike the reform concerning any single country that we summarized in the previous section, the 'multinational' and the 'international' reforms that we will discuss in this section have been partially influenced by two plans presented in the early 1940s by Keynes and Schumacher. As a matter of fact, together with but separately from Keynes, Schumacher was among the first proponents of an international settlement system using bank money instead of a commodity, like gold, to settle foreign trade.

The Schumacher (1943a, 1943b) plan aimed at setting up a multilateral clearing system between participating countries, to internalize all international payments for both paying and receiving countries. In such a system each country has a national clearing authority, so that

[t]he importer in country *A* pays for the goods he buys from country *B* by handing over to the Clearing Authority in his own country a sum of *A*-money which is deemed to discharge his debt. The exporter in country *B* receives from the Clearing Authority in his country an equivalent sum of *B*-money which is deemed to satisfy his claim.

(Schumacher 1943a: 150)

In such a multilateral clearing system, any international transaction has to be finally paid in domestic currency within the countries concerned and in some international monetary unit (to be issued) between them.

Keynes, whose plan was eventually presented at the 1944 Bretton Woods conference, elaborated upon this framework. He notably proposed to set up an ICU whereby all trade deficits are settled using an international bank money (bancor). In fact, Keynes linked the international bank money to gold, as in his plan the bancor 'would be defined in terms of a weight of gold' (Keynes 1980: 85). Further, in Keynes's plan gold had still some bearing on international settlements, as the ICU member countries were entitled to obtain a credit denominated in bancor by transferring their gold ownership to the ICU for the credit of their clearing account (*ibid.*: 175). According to Keynes, this is enough to

turn a stone (gold) into bread (ibid.: 177), in order to support economic growth (ibid.: 176).

Essentially, the Keynes plan aimed at reproducing between countries the monetary order that exists within any country, owing to the working of a payment system headed by the national central bank. As Keynes (1980: 44) explained in this regard,

[t]he idea underlying my proposal for a Currency Union is simple, namely to generalise the essential principle of banking, as it is exhibited within any closed system, through the establishment of an International Clearing Bank. This principle is the necessary equality of credits and debits, of assets and liabilities.

In reproducing at international level 'the essential principle of banking', Keynes went as far as to advocate that the International Clearing Bank (ICB) issue its means of payment for any transaction that it settles.

If no credits can be removed outside the banking system but only transferred within it, the [International Clearing] Bank *itself* can never be in difficulties. It can with safety make what advances it wishes to any of its customers with the assurance that the proceeds can only be transferred to the bank account of another customer.

(Keynes 1980: 44)

In this framework, the loans-generate-deposits mechanism (as explained by endogenous-money theorists like Moore (1988)) would imply that a deficit country (A) obtains from the ICB the amount of bancor that it needs in order to settle its foreign deficit against the relevant surplus country (B). As a result, the ICB would issue the bancor with a positive purchasing power: ICU member countries would be entitled to pay for their net imports by a stroke of the ICB pen, provided that their deficit does not exceed a predetermined quota (see Keynes 1980: 173). In such a system, Keynes is therefore obliged to provide a series of fines to prevent the piling up of credit and debit balances without limit in the accounts that countries hold at the ICB (ibid.: 173–5). As noted by Schumacher (1943b: 14), however,

the fines which under the [Keynes] Plan are to be imposed upon surplus countries must be viewed with a certain amount of doubt. They may help to achieve [balance-of-payments] equilibrium, but will they exert an expansionist pressure? Not if the surplus countries react to

them by cutting their exports, for then the result will be restriction and a contraction of world trade.

Further, '[i]f it appears doubtful whether the treatment proposed for excessive surpluses will, in fact, lead to an expansion of world trade, it is almost certain that the treatment proposed for excessive debits will be a restrictive force' (ibid.: 15). At another level, Schumacher argued that it would not 'be wise to place much reliance upon the efficacy of persuasive efforts emanating from an international authority' (ibid.: 14). The regulations proposed in Keynes's plan are thus problematic on two counts, granted

(1) that the suggested pressure by means mainly of fines and persuasion is quantitatively and qualitatively too weak to exert any measurable influence; and (2) that the expansionist effect which a penalization of surpluses might have (on special assumptions) may easily be neutralized, if not overcompensated, by the restrictive effect of the regulations applying to deficit countries.

(Schumacher 1943b: 15)

These institutional problems induced by the Keynes plan are aggravated by the fact that the creation of *bancor* generates a stock of money with no corresponding output, so that a country is not finally (that is, really) paid until it spends its *bancor* deposit and obtains real goods and services in exchange for it. This confusion between money and credit in the working of the ICB within the ICU is an essential weakness of Keynes's plan, which must be repaired to set up a reform capable to re-establish international monetary order as a result of an international system for final payments. Let us expand on this.

The 'multinational' reform

In this section we will analyse what would happen if a group of countries like those of the European Union, were to adopt a reform allowing for the final payment of their foreign transactions.

Let us assume that the European countries implementing the reform advocated here maintain (as the United Kingdom or Denmark) or recover their monetary sovereignty. Note that it is irrelevant whether this is done by reintroducing in most countries their original national currency or by letting them go on using their national euros, which would imply that the relevant countries understand that the euro is still far from being the unique currency of the euro area. It is in this

framework (which differs from the one we are actually confronted with only because euro-area member countries have given up their monetary sovereignty) that we analyse what we have called the 'multinational' reform.

This time the solution advocated concerns a group of countries and it is therefore possible to found it on the creation of a true and proper system of international payments, even if for a limited number of countries only. Each country participating in the new system has to create its own sovereign Bureau but, unlike the reform proposed in the previous section, no sovereign Bureau has to borrow and lend any sum of foreign currency abroad. The objective to provide deficit countries, for free, with the foreign currency they need to convey their real payments is no longer attained through a loan in foreign currency granted by their sovereign Bureaux, but through the institution of a supranational bank. In our example, this role is quite naturally assumed by the ECB, which is charged to issue an international, vehicular currency, say the international euro, €_i , to be used only as a *means* (as opposed to the object) for final payments between countries, both those members of the EU and those of the rest of the world, R. As in the Keynes plan, the €_i does not circulate within any given country and is issued according to the principle of double-entry bookkeeping. This new international currency has no intrinsic value and its function is solely that of conveying *real* payments between countries, without adding any monetary cost to them. In other words, the reform is conceived so that countries implementing it are no longer forced to purchase, at a positive cost equal to the sum purchased, the foreign currency required to convey the payment of their net imports.

In its new role, the ECB acts as an intermediary. It issues the €_i but under no circumstances does it finance any net purchase through an emission of €_i . The mechanism implemented by the ECB is that of the necessary equality between each country's sales and purchases. It is by complying with this that the ECB grants the vehicular or circular use of the €_i . All payments are carried out by countries' residents and the ECB merely conveys them in the international 'space'. If a country's overall purchases or imports exceed their sales or exports, it covers its deficit by transferring abroad the ownership over an equivalent part of its domestic production, and not by a net payment of €_i . Conversely, any exporting country is not paid in €_i , but in real goods (commercial or financial items), that is, by means of the vehicular use of the €_i . As a matter of fact money is by its very nature a numerical form with no intrinsic value whatsoever and can therefore not be the object of any

final payment, either nationally or internationally. The fact that the €_i is issued by the ECB does not change money's nature: the €_i is as valueless as any other money, and if a country were finally paid in €_i it would in reality accept a mere acknowledgment of debt in exchange for part of its domestic production.

In this regard, the only case that requires to be analysed in some details is that of the payment of a country's net total imports. Let us first assume that any given country, A, runs a deficit with respect to other EU countries, RE (the rest of Europe). To simplify the analysis, we propose again the numerical example used by Schmitt (2014) in his proposal for a reform, which can be adopted by any single country, and we assume that the exchange rate between A's domestic currency, MA, and the national currencies of RE (which we suppose to be a single country), and between these two currencies and the €_i is equal to $1 \text{ MA} = 1 \text{ MRE} = 1 \text{ €}_i$. If A's deficit is of $4 \text{ MA} = 4 \text{ MRE} = 4 \text{ €}_i$, country A's sovereign Bureau must ask the ECB to carry out its foreign payment to RE on its behalf. As a matter of fact, the payment will start at the pole of A's importers, residents of country A, who pay four MA to A's sovereign Bureau. Yet, the problem concerns the possibility for A to pay 4 €_i to RE in a situation where RE's purchases seem to fall short by 4 €_i . If the ECB were merely to pay 4 €_i to R on behalf of A (as in the Keynes plan), the reform would fail to its intentions and lead either to the non-payment of RE or to the double payment by A. It is in order to avoid such failure that the ECB has to act as a settlement institution and carry out its payments in €_i according to the real-time gross settlement protocol implemented domestically by national central banks. In our bilateral example and on the assumption that neither A nor RE has yet formed any positive deposit with the ECB, the payment of 4 €_i to RE has to be compensated by an equivalent payment to the benefit of A. This means that the reform must grant the reciprocity of A's and RE's payments even when A's purchases exceed its sales.

In the case under investigation, A's sovereign Bureau, which is a resident of country A, has to compensate the net imports of its other fellow residents by transferring to RE the ownership over a domestic output of A worth four MA. This result is no longer obtained, as in the case of a single country's reform, through a loan of foreign currency granted to R (or RE). A's sovereign Bureau does not borrow abroad any sum of foreign currency, since its foreign payments are carried out in €_i by the ECB. The Bureau of country A does also not borrow €_i from the ECB, whose function is not to *finance* countries' external purchases but merely to convey countries' reciprocal real payments. Thanks to the

intermediation of the ECB, moreover, no payment in MA carried out by A's residents reaches RE, which is paid in €_i by the ECB. It follows that the totality of A's domestic income is owned by A's residents, A's sovereign Bureau included, so that no foreign loan is needed to restore A's ownership over it. In particular, A's Bureau owns the four units of MA corresponding to A's net overall expenditures carried out in domestic currency by its residents. Now, this net gain in MA can be invested by the Bureau in A's domestic economy, and the Bureau can then sell abroad the claims on an equivalent part of A's national output. The sale of these financial claims by A's Bureau restores the balance-of-payments equilibrium between A and RE, and completes the ECB's intervention as a monetary intermediary and as a settlement institution.

The reciprocity of A's and RE's foreign transactions impedes the formation of A's external or sovereign debt and guarantees the respect of the balance-of-payments identity without reducing the level of employment in A. In Schmitt's 2014 (single country) reform, the level of employment is restored through the investment of the net gain of A's Bureau in a new production, additional to that of the rest of A's domestic economy, which loses part of its actual output to the benefit of R. In a 'multinational' or international reform, this loss no longer occurs and the gain of four MA is invested in A's current production. The final results, however, are similar: the level of employment is maintained, the country's external debt does not increase, and its foreign transactions are entirely and finally paid.

Once implemented by EU countries, the reform will also provide a new mechanism to convey the payments between these countries and the rest of the world. This would allow EU member countries to avoid the double payment of their net overall imports (if any) from R. Waiting for R to adopt and implement a reform similar to theirs, RE countries would be confronted with an alternative:

1. Ask the ECB to act as their collective sovereign Bureau *vis-à-vis* the rest of the world.
2. Ask each of their sovereign Bureaux to take charge of their payments outside the EU according to the principles of Schmitt's 2014 reform analysed above.

The first alternative (see point 1) requires a strong integration between EU countries and could thereby represent an important step toward a true monetary unification that would transform sovereign countries into regions of a single monetary area. Indeed, the net gain in domestic

currencies realized by the ECB in its function as common sovereign Bureau of EU member countries could be attributed to the single countries that generate it, but it could also be centralized and redistributed according to principles aiming at increasing the economic (and political) interdependence among EU countries. The second alternative (see point 2) is the one advocated by Schmitt for any single and sovereign country and does not entail any particular role for the ECB. Both will work fine and the choice between them will likely be influenced by the relevance attributed to the process of European monetary (and political) unification.

The world reform

A global solution to the problem of international payments requires the institution of a supranational bank, SB, whose role is to provide national central banks or national Bureaux with the international currency required to convey payments between countries. This international currency will have to be issued, free of cost, as a mere numerical vehicle enabling the exchange between commercial and financial goods defining countries' domestic resources. The institution of national Bureaux remains a necessary step, their role being once again that of being paid by national importers and pay national exporters in domestic money, and, additionally, of being credited and debited, in the international currency, by the SB. Like in the 'single country' and 'multinational' solutions, the goal will be to implement the balance of payments identity even when residents' total purchases or imports are initially greater than their total sales, and without charging the deficit country with an additional monetary payment. As established by monetary macroeconomics, payments are both real and monetary, the transfer to the payee of a given amount of real resources being conveyed through the 'vehicular' use of money. What a sound system of international payments prevents is the addition of the monetary to the real payment. This is achieved through the monetary intermediation of the SB, which will also be charged to act as international clearing house, allowing for the final settlement, in real time and on a gross basis, of every international transaction. A real-time gross settlement system managed by the SB, together with its international monetary intermediation, are the two crucial elements characterizing the international or world solution. The latter provides countries, free of cost, with the monetary means required to convey their payments internationally, while the former guarantees the final payment of international transactions by implementing the balance of payments

principle according to which every purchase is immediately matched by an equivalent sale.

The world solution can either be implemented by sovereign countries or by groups of countries having already adopted a multinational system of external payments. In the first case, each country will have to set up his own national Bureau, which will have to ask the SB to convey each one of its foreign payments. In other words, the reform on a world scale will substantially be the same as the one previously described with reference to a possible European level reform, the only difference being that if the reform is global, every country on the planet will be submitted to it and all the payments between countries will be carried out through the circular flow of a unique, supranational or international money issued by the SB. In the second case, the world reform will be the last step of a series of multinational reforms leading to the creation of a reduced number of monetary areas, each of which will regroup countries sharing a common, economic and political goal. A pyramidal structure, with the SB as its vertex, will guarantee the creation of a common monetary 'space' within which international payments between different monetary areas will be conveyed by an international or supranational currency issued at zero cost by the supranational bank.

Keynes's two main intuitions regarding his plan for a new system of international payments form also the core of Schmitt's proposals for a world reform: the use of an international currency issued by a supranational bank according to the principle of double-entry book-keeping as implemented by the national banking system, and the implementation of a multilateral clearing adopting a real-time gross final settlement mechanism. What in Keynes remained a mere intuition, acquires its full expression thanks to the crucial distinction between nominal money, a pure numerical vehicle, and real money, the result of the association between nominal money and real goods and services. As early as 1973, Schmitt was able to apply this distinction to the field of international payments in order to propose a plan of reform able to overcome the shortcomings still present in the Keynes plan, in particular its incapability to provide a purely numerical means to convey real payments between countries.

The world monetary reform advocated by Schmitt in 1973 is still valid. The major improvements it benefited from concern the plan of reform that can be implemented by a single country alone, in that it offers the possibility to reach a universal reform gradually. The institution of national Bureaux is another improvement that can be applied also to the world reform. Supranational bank and national Bureaux are

the only two institutional novelties introduced by the world reform, which permit international exchanges to be carried out without any additional monetary cost for countries still in need to import more than they export. Surpluses and deficits would still be possible under the aegis of the new system of international payments, but they would no longer carry with them the formation of a sovereign debt for the deficit countries, which would be protected against the pathological loss of their domestic resources. Surplus countries would be paid for the total of their exports, while deficit countries would not run into any sovereign debt. The surplus (difference between the sum paid by domestic importers and the sum paid to exporters) obtained by the Bureaux of deficit countries would be invested domestically and surplus countries would become the owners of part of the deficit countries' domestic output.

Put in a nutshell, these are the benefits to be obtained by the implementation of the reform in a world where countries still exist as sovereign entities and where international transactions are an important way to increase wealth. Whether this new world monetary order will be built all at once, through a new Bretton Woods-like agreement, or in successive steps, passing through reforms adopted by groups of countries, is not of fundamental relevance. What really matters is that the reasons for a reform, either for a 'single country' or a 'world' wide reform, are fully recognized and that the old-fashioned and cruelly mistaken non-system of international payments is finally abandoned and replaced by a system that conforms to the bookkeeping nature of money and to the logic of the macroeconomic laws it is based on.

Bibliography

- Bank for International Settlements (ed.) (2001) 'Marrying the macro- and micro-prudential dimensions of financial stability', *Bank for International Settlements Papers*, No. 1.
- Bell, S. (2001) 'The role of the state and the hierarchy of money', *Cambridge Journal of Economics*, 25 (2): 149–63.
- Bernanke, B.S. (1995) 'The macroeconomics of the Great Depression: A comparative approach', *Journal of Money, Credit and Banking*, 27 (1): 1–28.
- Bernstein, P.L. (1965) *A Primer on Money, Banking, and Gold*, New York: Vintage Books.
- Besomi, D. (2005) 'Clément Juglar and the transition from crises theory to business cycle theories', Paper prepared for a conference on the occasion of the centenary of the death of Clément Juglar, Paris, 2 December.
- Blanchard, O.J. (2000) 'What do we know about macroeconomics that Fisher and Wicksell did not?' *National Bureau of Economic Research Working Papers*, No. 7550.
- Blanchard, O.J. (2009) 'The state of macro', *The Annual Review of Economics*, 1: 209–28.
- Böhm-Bawerk, E. (1898) 'Review of Eugen von Bergmann *Die Wirtschaftskrisen: Geschichte der Nationalökonomischen Krisentheorien*', *Zeitschrift für Volkswirtschaft, Sozialpolitik und Verwaltung*, 7: 132–3.
- Böhm-Bawerk, E. (1884/1959) *Capital and Interest*, Vol. I *History and Critique of Interest Theories*, South Holland: Libertarian Press.
- Böhm-Bawerk, E. (1889/1959) *Capital and Interest*, Vol. II *Positive Theory of Capital*, South Holland: Libertarian Press.
- Böhm-Bawerk, E. (1909/1959) *Capital and Interest*, Vol. III *Further Essays on Capital and Interest*, South Holland: Libertarian Press.
- Borio, C. and Lowe, P. (2002) 'Asset prices, financial and monetary stability: Exploring the nexus', *Bank for International Settlements Working Papers*, No. 114.
- Borio, C. and White, W. (2004) 'Whither monetary and financial stability? The implications of evolving policy regimes', *Bank for International Settlements Working Papers*, No. 147.
- Cencini, A. (1982) 'The logical indeterminacy of relative prices', in M. Baranzini (ed.) *Advances in Economic Theory*, Oxford: Blackwell, pp. 126–36.
- Cencini, A. (1984) *Time and the Macroeconomic Analysis of Income*, London and New York: Pinter Publishers and St. Martin's Press (republished by Bloomsbury 2013).
- Cencini, A. (1988) *Money, Income and Time*, London and New York: Pinter Publishers (republished by Bloomsbury 2013).
- Cencini, A. (1996) 'Inflation and deflation: The two faces of the same reality', in A. Cencini and M. Baranzini (eds) *Inflation and Unemployment. Contributions to a New Macroeconomic Approach*, London and New York: Routledge, pp. 17–60.
- Cencini, A. (1997) *Monetary Theory, National and International*, London and New York: Routledge.

- Cencini, A. (2001) *Monetary Macroeconomics; A New Approach*, London and New York: Routledge.
- Cencini, A. (2003a) 'IS-LM: A final rejection', in L.-P. Rochon and S. Rossi (eds) *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 295–321.
- Cencini, A. (2003b) 'Micro, macro et l'analyse du circuit', in P. Piégay and L.-P. Rochon (eds) *Théories monétaires post Keynésiennes*, Paris: Economica, pp. 209–25.
- Cencini, A. (2005) *Macroeconomic Foundations of Macroeconomics*, London and New York: Routledge.
- Cencini, A. (2008) *Elementi di macroeconomia monetaria*, Milan: CEDAM.
- Cencini, A. (2012a) 'Towards a macroeconomic approach to macroeconomics', in C. Gnos and S. Rossi (eds) *Modern Monetary Macroeconomics; A New Paradigm for Economic Policy*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 39–68.
- Cencini, A. (2012b) 'Is there a common cause to economic and financial crises?' in C. Gnos and S. Rossi (eds) *Modern Monetary Macroeconomics; A New Paradigm for Economic Policy*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 193–217.
- Cencini, A. and Baranzini, M. (eds) (1996) *Inflation and Unemployment; Contributions to a New Macroeconomic Approach*, London and New York: Routledge.
- Cencini, A. and Schmitt, B. (1991) *External Debt Servicing; A Vicious Circle*, London and New York: Pinter Publishers.
- Champernowne, D.G. (1936) 'Unemployment, basic and monetary: The classical analysis and the Keynesian', *Review of Economic Studies*, 3 (3): 201–16.
- Chari, V.V. and Kehoe, P.J. (2007) 'The heterogeneous state of modern macroeconomics: A reply to Solow', *Federal Reserve Bank of Minneapolis Research Department Staff Reports*, No. 399.
- Chick, V. (1983) *Macroeconomics after Keynes: A Reconsideration of the General Theory*, Oxford: Philip Allan.
- Chick, V. (2000) 'Money and effective demand', in J. Smithin (ed.) *What Is Money?* London and New York: Routledge, pp. 124–38.
- Coddington, A. (1976) 'Keynesian economics, the search for first principles', *Journal of Political Literature*, 14 (4): 1258–73.
- Coddington, A. (1983) *Keynesian Economics: The Search for First Principles*, London: Allen & Unwin.
- D'Arista, J. (2009) 'The evolving international monetary system', *Cambridge Journal of Economics*, 33 (4): 633–52.
- D'Arista, J. and Ertürk, K.A. (2010) 'The case for international monetary reform', *Real World Economics Review*, 55 (4): 58–81.
- Dal-Pont Legrand, M. and Hagemann, H. (2007) 'Business cycles in Juglar and Schumpeter', *History of Economic Thought*, 49 (1): 1–17.
- Davidson, P. (1978) *Money and the Real World*, London: Macmillan, second edition (first published 1972).
- Davidson, P. (1991) 'Three views on inflation: Monetarist, neoclassical Keynesian and post Keynesian', in *Controversies in Post Keynesian Economics*, Aldershot and Brookfield: Edward Elgar, pp. 83–110.

- Davidson, P. (1998) 'Post Keynesian employment analysis and the macroeconomics of OECD unemployment', *Economic Journal*, 108 (448): 817–31.
- Debreu, G. (1959) *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*, New Haven: Yale University Press.
- Diamond, D.W. and Rajan, R.G. (2001) 'Banks, short-term debt and financial crises: Theory, policy implications and applications', *Carnegie-Rochester Conferences Series on Public Policy*, 54: 37–71.
- Dimand, R.W. (1988) 'Review of *Interpreting Mr. Keynes: The IS-LM Enigma* by Warren Young', *Canadian Journal of Economics*, 21 (3): 659–62.
- Dooley, M.P., Folkerts-Landau, D. and Garber, P.M. (2009) 'Bretton Woods II still defines the international monetary system', *National Bureau of Economic Research Working Papers*, No. 14731.
- Dymski, G.A. (2011) 'Keynesian approaches to financial crisis', in E. Hein and E. Stockhammer (eds) *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 325–51.
- Epstein, G.A. (ed.) (2005) *Financialization and the World Economy*, Cheltenham, UK and Northampton, MA: Edward Elgar.
- Fetter, F.A. (1914) 'Interest theories, old and new', *American Economic Review*, 4 (1): 68–92.
- Fisher, I. (1907) *The Rate of Interest: Its Nature, Determination and Relation to Economic Phenomena*, New York: Macmillan.
- Fisher, I. (1911) *The Purchasing Power of Money*, New York: Macmillan.
- Fisher, I. (1913) 'The impatience theory of interest', *American Economic Review*, 3 (3): 610–18.
- Fisher, I. (1933) 'The debt-deflation theory of great depressions', *Econometrica*, 1 (4): 337–57.
- Fischer, S. (1983) 'A framework for monetary and banking analysis', *Economic Journal*, 93 (conference supplement): 1–16.
- Fontana, G. (2009) *Money, Uncertainty and Time*, Abingdon and New York: Routledge.
- Friedman, M. (ed.) (1956) *Studies in the Quantity Theory of Money*, Chicago: Chicago University Press.
- Friedman, M. (1970) 'A theoretical framework for monetary analysis', *Journal of Political Economy*, 78 (2): 193–238.
- Friedman, M. and Schwartz, A.J. (1963) *A Monetary History of the United States: 1867–1960*, Princeton: Princeton University Press.
- Gali, J. (2008) *Monetary Policy, Inflation, and the Business Cycle*, Princeton and Oxford: Princeton University Press.
- Gali, J. (2012) 'Notes for a new guide to Keynes (I): Wages, aggregate demand, and employment', *National Bureau of Economic Research Working Papers*, No. 18651.
- Gali, J., Gertler, M. and López-Salido, J.D. (2007) 'Markups, gaps, and the welfare costs of business fluctuations', *Review of Economics and Statistics*, 89 (1): 44–59.
- Garegnani, P. (1972) *Il capitale nelle teorie della distribuzione*, Milano: Giuffrè Editore.
- Gnos, C. (1998) 'The Keynesian identity of income and output', in P. Fontaine and A. Jolink (eds) *Historical Perspectives on Macroeconomics: Sixty Years after the General Theory*, London and New York: Routledge, pp. 40–8.

- Goodfriend, M. and King, R.G. (1997) 'The new neoclassical synthesis and the role of monetary policy', in B.S. Bernanke and J. Rotemberg (eds) *NBER Macroeconomics Annual 1997*, Vol. 12, Harvard: MIT Press, pp. 231–96.
- Goodhart, C.A.E. (1989) *Money, Information and Uncertainty*, London and Basingstoke: Macmillan, second edition (first published 1975).
- Goodhart, C.A.E. (1998) 'The two concepts of money: Implications for the analysis of optimal currency areas', *European Journal of Political Economy*, 14 (3): 407–32.
- Goodhart, C.A.E. (2007) 'Whatever became of the monetary aggregates?' Peston Lecture in honour of Maurice, Lord Peston, delivered at Queen Mary College, London, 28 February.
- Gordon, R.J. (1990) 'What is new Keynesian economics?' *Journal of Economic Literature*, 28 (3): 1115–71.
- Gould, S.J. (2007) *The Richness of Life*, London: Vintage Books.
- Graziani, A. (1990) 'The theory of the monetary circuit', *Economies et Sociétés*, 24 (6): 7–36.
- Graziani, A. (2003) *The Monetary Theory of Production*, Cambridge: Cambridge University Press.
- Greenwald, B. and Stiglitz, J. (1993) 'New and old Keynesians', *Journal of Economic Perspectives*, 7 (1): 23–44.
- Hansen, A.H. (1953) *A Guide to Keynes*, New York: McGraw Hill.
- Harrod, R.F. (1937) 'Keynes and traditional theory', *Econometrica*, 5 (1): 74–86.
- Hayek, F.A. (1933/2008) 'Monetary theory and the trade cycle', *Ludwig von Mises Institute*, available at <http://mises.org/articles.aspx?AuthorId=126>: 1–52 (first published 1933).
- Hicks, J.R. (1933/1982) 'Equilibrium and the cycle', in *Money, Interest and Wages: Collected Essays on Economic Theory*, Oxford: Basil Blackwell, Vol. II, pp. 28–41.
- Hicks, J.R. (1937/1982) 'Mr. Keynes and the "Classics": A suggested interpretation', in *Money, Interest and Wages: Collected Essays on Economic Theory*, Oxford: Basil Blackwell, Vol. II, pp. 100–15.
- Hicks, J.R. (1965) *Capital and Growth*, Oxford: Oxford University Press.
- Hicks, J.R. (1976/1982) 'Time in economics', in *Money, Interest and Wages: Collected Essays on Economic Theory*, Oxford: Basil Blackwell, Vol. II, pp. 282–300.
- Hicks, J.R. (1980–81/1982) 'IS–LM: An explanation', in *Money, Interest and Wages: Collected Essays on Economic Theory*, Oxford: Basil Blackwell, Vol. II, pp. 318–31.
- Hülsmann, J.G. (2002) 'A theory of interest', *Quarterly Journal of Austrian Economics*, 5 (4): 77–110.
- Hume, D. (1826) 'The Philosophical Works', in Rotwein, E. (ed.) *David Hume: Writings on Economics*, Edinburgh: T. Nelson & Sons, Vol. III, reprinted.
- Hutt, W.H. (1974) *A Rehabilitation of Say's Law*, Athens, OH: Ohio University Press.
- Ingham, G. (1996) 'Money is a social relation', *Review of Social Economy*, 54 (4): 507–29.
- Jevons, W.S. (1884) *Investigations in Currency and Finance*, edited by H.S. Foxwell, London: Macmillan.
- Juglar, C. (1857) 'Des crises commerciales et monétaires de 1800 à 1857', *Journal des Economistes*, 14 (April and May): 35–60, 255–67.
- Juglar, C. (1889) *Des crises commerciales et de leur retour périodique en France, en Angleterre et aux Etats-Unis*, Paris: Alcan, second edition (reprinted New York: Kelly, 1967).

- Kadmos, G.A. and O'Hara, P.A. (2000) 'The taxes-drive-money and employer of last resort approach to government policy', *Journal of Economic and Social Policy*, 5 (1): 1–22.
- Kaldor, N. (1983) 'Keynesian economics after fifty years', in D. Worswick and J. Trevithick (eds) *Keynes and the Modern World*, Cambridge: Cambridge University Press, pp. 1–28.
- Keen, S. (2011) 'Debunking macroeconomics', *Economic Analysis and Policy*, 41 (3): 147–67.
- Keen, S. (2013) 'Predicting the "global financial crisis": Post-Keynesian macroeconomics', *Economic Record*, 89 (285): 228–54.
- Keynes, J.M. (1930/1971) *A Treatise on Money*, Vol. I: *The Pure Theory of Money*, Cambridge: Macmillan and St Martin's Press.
- Keynes, J.M. (1936/1946) *The General Theory of Employment, Interest and Money*, London and Basingstoke: Macmillan.
- Keynes, J.M. (1973a) *The Collected Writings of John Maynard Keynes*, Vol. XIII: *The General Theory and After: Part I. Preparation*, London and Basingstoke: Macmillan.
- Keynes, J.M. (1973b) *The Collected Writings of John Maynard Keynes*, Vol. XIV: *The General Theory and After: Part II. Defence and Development*, London and Basingstoke: Macmillan.
- Keynes, J.M. (1980) *The Collected Writings of John Maynard Keynes*, Vol. XXV: *Activities 1940–1944. Shaping the Post-War World: The Clearing Union*, London and Basingstoke: Macmillan.
- King, J.E. (2012) *The Microfoundations Delusion: Metaphor and Dogma in the History of Macroeconomics*, Cheltenham, UK and Northampton, MA: Edward Elgar.
- King, R.G. and Plosser, C. (1984) 'Money, credit and prices in a real business cycle', *American Economic Review*, 74 (3): 363–80.
- Klein, L.R. (1947) *The Keynesian Revolution*, New York: Macmillan.
- Kregel, J. (2008) 'Observations on the problem of "too big to fail/save/resolve"', *Levy Economics Institute of Bard College Public Policy Note*, No. 11.
- Kregel, J. (2009) 'Minsky's cushions of safety: Systemic risk and the crisis in the U.S. subprime mortgage market', *Levy Economics Institute of Bard College Public Policy Brief*, No. 93.
- Krugman, P. (1999) 'The return of depression economics', *Foreign Affairs*, 78 (1): 56–74.
- Lawson, J.A. (1848) 'The cause of commercial panics', *Bankers' Magazine*, 8: 415–20.
- Lucas, R.E., Jr. (1972) 'Expectations and the neutrality of money', *Journal of Economic Theory*, 4 (2): 103–24.
- Lucas, R.E., Jr. (1977) 'Understanding business cycles', in K. Brunner and A.H. Meltzer (eds) *Stabilization of the Domestic and International Economy*, Amsterdam: North-Holland, pp. 7–29.
- Lucas, R.E., Jr. (1980) 'Methods and problems in business cycle theory', *Journal of Money, Credit and Banking*, 12 (4): 696–715.
- Lucas, R.E., Jr. (1981) *Studies in Business-Cycle Theory*, Oxford: Basil Blackwell.
- Lucas, R.E., Jr. (2003) 'Macroeconomic priorities', *American Economic Review*, 93 (1): 1–14.
- Lucas, R.E., Jr. and Sargent, T.J. (1978) 'After Keynesian macroeconomics', *Journal of Political Economy*, 86 (1): 71–86.

- Luxemburg, R. (1913) *Die Akkumulation des Kapitals: Ein Beitrag zur ökonomischen Erklärung des Imperialismus*, Berlin: Paul Singer.
- Mankiw, N.G. (2006) 'The macroeconomist as scientist and engineer', *Journal of Economics Perspectives*, 20 (4): 29–46.
- Mankiw, N.G. (2007) *Macroeconomics*, New York: Worth Publishers.
- Marx, K. (1839/1973) *Grundrisse*, Harmondsworth: Penguin.
- Marx, K. (1867/1976) *Capital, Vol. I*, Harmondsworth: Penguin.
- Marx, K. (1885/1978) *Capital, Vol. II*, Harmondsworth: Penguin.
- Marx, K. (1894/1981) *Capital, Vol. III*, Harmondsworth: Penguin.
- Meade, J.E. (1937) 'A simplified model of Mr Keynes' system', *Review of Economic Studies*, 4 (1): 98–107.
- Mills, J. (1868) 'On credit cycles and the origin of commercial panics', *Transactions of the Manchester Statistical Society*, Session 1867–8: 5–40.
- Minsky, H.P. (1957a) 'Central banking and money market changes', *Quarterly Journal of Economics*, 71 (2): 171–87.
- Minsky, H.P. (1957b) 'Monetary systems and accelerator models', *American Economic Review*, 47 (6): 860–83.
- Minsky, H.P. (1975) *John Maynard Keynes*, New York: Columbia University Press.
- Minsky, H.P. (1977) 'The financial instability hypothesis: An interpretation of Keynes and an alternative to "standard" theory', *Nebraska Journal of Economics and Business*, 16 (1): 5–16.
- Minsky, H.P. (1982) *Can "It" Happen Again? Essays on Instability and Finance*, Armonk, NY: M.E. Sharpe.
- Minsky, H.P. (1986) *Stabilizing an Unstable Economy*, New Haven: Yale University Press.
- Minsky, H.P. (1992) 'Profits, deficits and economic instability: A policy discussion', in D.B. Papadimitriou (ed.) *Profits, Deficits and Instability*, London and Basingstoke: Macmillan, pp. 11–22.
- Mishkin, F.S. (2004) *The Economics of Money, Banking and Financial Markets*, New York: Pearson and Addison Wesley.
- Mitchell Innes, A. (1913) 'What is money?' *Banking Law Journal*, May: 377–408.
- Modigliani, F. (1944) 'Liquidity preference and the theory of interest and money', *Econometrica*, 12 (1): 45–88.
- Moore, B.J. (1988) *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge: Cambridge University Press.
- Muth, J.F. (1961) 'Rational expectations and the theory of price movements', *Econometrica*, 29 (3): 315–35.
- Nersisyan, Y. and Wray, L.R. (2010) 'The global financial crisis and the shift to shadow banking', *European Journal of Economics and Economic Policy: Intervention*, 7 (2): 377–400.
- Ormerod, P. (1994) *The Death of Economics*, London: Faber and Faber.
- Parguez, A. and Seccareccia, M. (2000) 'The credit theory of money: The monetary circuit approach', in J. Smithin (ed.) *What Is Money?* London and New York: Routledge, pp. 101–23.
- Plosser, C.I. (1989) 'Understanding real business cycles', *Journal of Economic Perspectives*, 3 (3): 51–77.
- Prescott, E.C. (1999) 'Some observations on the Great Depression', *Federal Reserve Bank of Minneapolis Quarterly Review*, 23 (1): 25–31.

- Quesnay, F. (1758/2005) *Œuvres économiques complètes et autres textes*, C. Théré, L. Charles and J.-C. Perrot (eds), Paris: Institut national d'études démographiques.
- Reddaway, W.B. (1936) 'The General Theory of Employment, Interest and Money by J.M. Keynes', *Economics Record*, 12: 28–36.
- Reinhart, C.M. and Rogoff, K.S. (2008) 'Is the 2007 U.S. subprime crisis so different? An international historical comparison', *American Economic Review*, 98 (2): 339–44.
- Reinhart, C.M. and Rogoff, K.S. (2009) 'The aftermath of financial crises', *American Economic Review: Papers & Proceedings*, 99 (2): 466–72.
- Ricardo, D. (1817/1951) *On the Principles of Political Economy and Taxation*, Cambridge: Cambridge University Press.
- Ricardo, D. (1824) *Plan for the Establishment of a National Bank*, in P. Sraffa and M.H. Dobb (eds) *The Works and Correspondence of David Ricardo*, Vol. IV: *Pamphlets and Papers 1815–1823*, Cambridge: Cambridge University Press, pp. 276–300.
- Ricardo, D. (1953) *The Works and Correspondence of David Ricardo*, Vol. VIII: *Letters 1819–June 1821*, Cambridge: Cambridge University Press.
- Robertson, D.H. (1937) *Money*, Cambridge: Cambridge University Press.
- Robinson, J. (1953–54) 'The production function and the theory of capital', *Review of Economic Studies*, 21 (2): 81–106.
- Robinson, J. (1962) 'Review of H.G. Johnson's *Money, Trade and Economic Growth*', *Economic Journal*, 72 (287): 690–2.
- Robinson, J. (1971) *Economics Heresies*, London: Macmillan.
- Rochon, L.-P. (1999) *Credit, Money and Production: An Alternative Post-Keynesian Approach*, Cheltenham, UK and Northampton, MA: Edward Elgar.
- Rochon, L.-P. and Rossi, S. (2003) 'Introduction', in L.-P. Rochon and S. Rossi (eds) *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. xx–lvi.
- Roelandts, M. (2009) 'Le cadre méthodologique de la théorie des crises chez Marx et sa validation empirique', available at www.capitalisme-et-crise.info/fr/.
- Romer, D. (1993) 'The new Keynesian synthesis', *Journal of Economic Perspectives*, 7 (1): 5–22.
- Rossi, S. (2001) *Money and Inflation: A New Macroeconomic Analysis*, Cheltenham, UK and Northampton, MA: Edward Elgar (reprinted 2003).
- Rossi, S. (2007a) *Money and Payments in Theory and Practice*, London and New York: Routledge.
- Rossi, S. (2007b) 'The monetary-policy relevance of an international settlement institution: The Keynes plan 60 years later', in A. Giacomini and M.C. Marcuzzo (eds) *Money and Markets: A Doctrinal Approach*, London and New York: Routledge, pp. 96–114.
- Rossi, S. (2008) 'The role of banks in the (over-)accumulation of capital', *European Journal of Economic and Social Systems*, 21 (2): 213–31.
- Rossi, S. (2009a) 'Taxes-drive-money approach', in P.A. O'Hara (ed.) *International Encyclopedia of Public Policy, Volume 2: Economic Policy*, Perth: GPERU, pp. 628–40.
- Rossi, S. (2009b) 'Monetary circuit theory and money emissions', in J.-F. Ponsot and S. Rossi (eds) *The Political Economy of Monetary Circuits: Tradition and Change in Post-Keynesian Economics*, Basingstoke and New York: Palgrave Macmillan, pp. 36–55.

- Rossi, S. (2010) 'Financial stability requires macroeconomic foundations of macroeconomics', *Journal of Philosophical Economics*, 3 (2): 58–73.
- Rossi, S. (2011a) 'Macro and financial economics need a quantum leap', *International Journal of Pluralism and Economics Education*, 2 (3): 306–17.
- Rossi, S. (2011b) 'Can it happen again? Structural policies to avert further systemic crises', *International Journal of Political Economy*, 40 (2): 61–78.
- Rossi, S. (2012) 'The monetary–structural origin of TARGET2 imbalances across Euroland', in C. Gnos and S. Rossi (eds) *Modern Monetary Macroeconomics: A New Paradigm for Economic Policy*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 221–38.
- Rossi, S. (2013) 'Structural reforms to reduce systemic financial fragility', in N. Karagiannis, Z. Madjd-Sadjadi and S. Sen (eds) *The US Economy and Neoliberalism: Alternative Strategies and Policies*, Abingdon and New York: Routledge, pp. 153–63.
- Rossi, S. (2014) 'A structural monetary reform to reduce global imbalances: Keynes's plan revisited to avert international payment deficits', in R. Bellofiore and G. Vertova (eds) *The Great Recession and the Contradictions of Contemporary Capitalism*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 134–50.
- Rowthorn, R.E. (1977) 'Conflict, inflation and money', *Cambridge Journal of Economics*, 1 (3): 215–39.
- Rueff, J. (1963) 'Gold exchange standard: A danger to the west', in H.G. Grubel (ed.) *World Monetary Reform: Plans and Issues*, Stanford and London: Stanford University Press and Oxford University Press, pp. 320–8.
- Samuelson, P.A. (1948) *Economics*, New York: McGraw Hill.
- Sawyer, M. (2003) 'Employer of last resort: Could it deliver full employment and price stability?' *Journal of Economic Issues*, 37 (4): 881–907.
- Say, J.B. (1803/1972) *Traité d'économie politique*, Paris: Calmann-Lévy.
- Say, J.B. (1821) *Treatise on Political Economy*, London: Longman, Hurst, Rees Orme, and Brown Paternoster-Row.
- Schmitt, B. (1966) *Monnaie, salaires et profits*, Paris: Presses Universitaires de France.
- Schmitt, B. (1972) *Macroeconomic Theory: A Fundamental Revision*, Albeuve: Castella.
- Schmitt, B. (1975) *Théorie unitaire de la monnaie, nationale et internationale*, Albeuve: Castella.
- Schmitt, B. (1984a) *Inflation, chômage et malformations du capital*, Paris and Albeuve: Economica and Castella.
- Schmitt, B. (1984b) *La France souveraine de sa monnaie*, Paris and Albeuve: Economica and Castella.
- Schmitt, B. (1985) 'Introduzione agli scritti monetari di David Ricardo', in D. Ricardo *Scritti monetari*, Rome: Istituto della Enciclopedia Italiana, pp. 3–85.
- Schmitt, B. (1993) *Teoria unitaria della moneta, nazionale e internazionale*, Naples: Liguori.
- Schmitt, B. (1993–94) *Notes sur la théorie de l'intérêt*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1995–96) *Cours de théorie monétaire*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1996a) 'A new paradigm for the determination of money prices', in G. Deleplace and E.J. Nell (eds) *Money in Motion: The Post Keynesian and*

- Circulation Approaches*, London and New York: Macmillan and St Martin's Press, pp. 104–38.
- Schmitt, B. (1996b) *Cours de théorie monétaire*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1997) *Note à l'attention des étudiants*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1998–99a) *Le chômage et son éradication*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1998–99b) *Annexe au cours écrit*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1999a) *Critique fondamentale de la pensée néoclassique*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1999b) *Preuve analytique de l'inanité de la pensée néoclassique*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (1999c) *En marge du cours: quelques considérations sur les nombres en économie, dans la pensée néoclassique et au-delà*, Fribourg: University of Fribourg, mimeo.
- Schmitt, B. (2000) 'The double charge of external debt servicing' *Quaderno di ricerca No. 1*, Lugano-Vezia: Research Laboratory in Monetary Economics.
- Schmitt, B. (2005) *Théorème de l'intérêt: le double poids des intérêts afférents aux dettes extérieures*, Lugano-Vezia: Research Laboratory in Monetary Economics, mimeo.
- Schmitt, B. (2007) *Théorème de l'intérêt*, Lugano-Vezia: Research Laboratory in Monetary Economics.
- Schmitt, B. (2012) 'Sovereign debt and interest payments', in C. Gnos and S. Rossi (eds) *Modern Monetary Macroeconomics: A New Paradigm for Economic Policy*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 239–60.
- Schmitt, B. (2014) *The Formation of Sovereign Debt: Diagnosis and Remedy*, SSRN.
- Schmitt, B. and De Gottardi, C. (2003) 'An internal critique of general equilibrium theory', in L.-P. Rochon and S. Rossi (eds) *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 265–94.
- Schumacher, E.F. (1943a) 'Multilateral clearing', *Economica*, 10 (38): 150–65.
- Schumacher, E.F. (1943b) 'The new currency plans', *Bulletin of the Oxford University Institute of Statistics*, 5 (Supplement): 8–28.
- Schumpeter, J.A. (1950) 'Wesley Clair Mitchell (1874–1948)', *Quarterly Journal of Economics*, 64 (1): 139–55.
- Schumpeter, J.A. (1954/1994) *History of Economic Analysis*, London: Allen & Unwin.
- Seager, H.R. (1912) 'The impatience theory of interest', *American Economic Review*, 2 (4): 834–51.
- Smets, F. and Wouters, R. (2007) 'Shocks and frictions in US business cycles: A Bayesian DSGE approach', *American Economic Review*, 97(3): 586–606.
- Smith, A. (1776/1976) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Oxford: Clarendon Press.
- Snowdon, B., Vane, H. and Wynarczyk, P. (1994) *A Modern Guide to Macroeconomics*, Cheltenham, UK and Northampton, MA: Edward Elgar.

- Sowell, T. (1972) *Say's Law*, Princeton: Princeton University Press.
- Sowell, T. (1994) *Classical Economics Reconsidered*, Princeton: Princeton University Press.
- Stamp, M. (1963) 'The Stamp plan – 1962 version', in H.G. Grubel (ed.) *World Monetary Reform: Plans and Issues*, Stanford and London: Stanford University Press and Oxford University Press, pp. 80–9.
- Stockhammer, E. (2011) 'The macroeconomics of unemployment', in E. Hein and E. Stockhammer (eds) *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 137–64.
- Tobin, J. (1963) 'Commercial banks as creators of "money"', in D. Carson (ed.) *Banking and Monetary Studies*, Homewood, IL: Richard D. Irwin, pp. 408–19.
- Tobin, J. (1965) 'Money and economic growth', *Econometrica*, 33 (4): 671–84.
- Tobin, J. and Golub, S.S. (1998) *Money, Credit and Capital*, Boston: Mc Graw-Hill.
- Turgot, A.R.J. (1767/2011) *The Turgot Collection*, edited by D. Gordon, Auburn, AL: Ludwig von Mises Institute.
- Von Mises, L. (1912/1981) *The Theory of Money and Credit*, Indianapolis: Liberty Classics.
- Von Mises, L. (1928) *Geldwertstabilisierung und Konjunkturpolitik*, Jena: Gustav Fischer.
- Walras, L. (1874/1984) *Elements of Pure Economics*, Philadelphia: Orion Editions.
- Wicksell, K. (1934) *Lectures on Political Economy*, London: Routledge.
- Wicksell, K. (1954) *Value, Capital and Rent*, London: Allen & Unwin.
- Wicksell, K. (1898/1965) *Interest and Prices*, New York: Kelly.
- Woodford, M. (1999) 'Revolution and evolution in twentieth-century macroeconomics', Paper presented at the conference on 'Frontiers of the Mind in the Twenty-First Century', Washington.
- Woodford, M. (2003) *Interest and Prices*, Princeton: Princeton University Press.
- Woodford, M. (2007) 'How important is money in the conduct of monetary policy?' *Columbia University Discussion Papers*, No. 0607–16.
- Woodford, M. (2008) 'Convergence in macroeconomics: Elements of the new synthesis', Paper presented at the annual meeting of the American Economic Association, New Orleans.
- Wray, L.R. (1998) *Understanding Modern Money: The Key to Full Employment and Price Stability*, Cheltenham, UK and Northampton, MA: Edward Elgar.
- Wray, L.R. (2003a) 'Seigniorage or sovereignty?' in L.-P. Rochon and S. Rossi (eds) *Modern Theories of Money: The Nature and Role of Money in Capitalist Economies*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 84–102.
- Wray, L.R. (2003b) 'The neo-chartalist approach to money', in S.A. Bell and E.J. Nell (eds) *The State, the Market and the Euro: Chartalism versus Metallism in the Theory of Money*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 89–110.
- Wray, L.R. (2008) 'Lessons from the subprime meltdown', *Challenge*, 51 (2): 40–68.
- Wray, L.R. (2009) 'The rise and fall of money manager capitalism: A Minskian approach', *Cambridge Journal of Economics*, 33 (4): 807–28.

- Wray, L.R. (2010) 'What should banks do? A Minsky analysis', *Levy Economics Institute of Bard College Public Policy Brief*, No. 115.
- Wray, L.R. (2012) *Modern Money Theory: A Primer on Macroeconomics for Sovereign Monetary Systems*, London and New York: Palgrave Macmillan.
- Young, W. (1987) *Interpreting Mr. Keynes: The IS-LM Enigma*, Cambridge: Polity Press.

Author Index

- Bell, S., 124
Bernanke, B.S., 72–3, 76
Bernstein, P.L., 30
Besomi, D., 59
Blanchard, O.J., 99–100, 117, 119–20, 134
Böhm-Bawerk, E., 10, 66, 78, 151, 158, 160–4, 167, 185–90, 198
Borio, C., 79
- Champernowne, D.G., 109
Chari, V.V., 119
Chick, V., 123–4
Coddington, A., 107
- Dal-Pont Legrand, M., 76, 78
D'Arista, J., 242
Davidson, P., 123, 127
Debreu, G., 74–5, 88, 98, 100
Diamond, D.W., 72
Dimand, R.W., 109
Dooley, M.P., 241
Dymski, G.A., 128
- Epstein, G.A., 78
Ertürk, K.A., 242
- Fetter, F.A., 185–90, 198
Fischer, S., 227
Fisher, I., 81, 84–6, 94, 189–90
Fontana, G., 124
Friedman, M., 66, 76, 83–7, 96
- Gali, J., 119, 134, 146
Garegnani, P., 162, 164
Gnos, C., 126
Golub, S.S., 21
Goodfriend, M., 96, 98, 102, 119, 145
Goodhart, C.A.E., 97, 123–4
Gordon, R.J., 118
Gould, S.J., 68
Graziani, A., 125–6
Greenwald, B., 110, 117–18
- Hagemann, H., 76, 78
Hansen, A.H., 109
Harrod, R.F., 109
Hayek, F.A., 35, 66–9, 76–8, 87
Hicks, J.R., 65–6, 106, 109–13, 116, 122
Hülsmann, J.G., 186–7
Hume, D., 94
Hutt, W.H., 39
- Ingham, G., 124
- Jevons, W.S., 59–60, 64
Juglar, C., 59–60, 64, 67, 76
- Kadmos, G.A., 125
Kahn, A.E., 34, 108
Kaldor, N., 41
Keen, S., 227–8, 239
Kehoe, P.J., 119
Keynes, J.M., 16–19, 22–4, 39, 41–3, 48, 50–3, 55, 78–9, 106–28, 132–4, 136, 139, 145, 147–8, 151, 153, 161, 165–7, 191–3, 198, 226, 232, 242, 251–5, 258
King, J.E., 128
King, R.G., 96, 98–9, 102, 119, 145
Klein, L.R., 109
Kregel, J., 81, 128
Krugman, P., 72
- Lawson, J.A., 59–60, 64
Lowe, P., 79
Lucas, R.E., Jr., 66–8, 71, 76, 95–8, 100, 118–19
Luxemburg, R., 63
- Mankiw, N.G., 17–18, 21, 104, 117
Marx, K., 10, 23, 29, 60–3, 130–1, 151, 156, 158–60, 163, 167, 175, 201
Meade, J.E., 109
Mills, J., 59, 64

- Minsky, H.P., 9, 60, 78–82, 124, 128, 203, 227
- Mishkin, F.S., 17–18, 21
- Mitchell Innes, A., 21
- Modigliani, F., 109
- Moore, B.J., 229, 252
- Muth, J.F., 66, 96, 98
- Nersisyan, Y., 81
- O'Hara, P.A., 125
- Ormerod, P., 128
- Parguez, A., 30–1, 33
- Plosser, C., 65–6, 77, 98–9
- Prescott, E.C., 65
- Quesnay, F., 152
- Rajan, R.G., 72
- Reddaway, W.B., 109
- Reinhart, C.M., 70
- Ricardo, D., 10–11, 29, 151, 156–9, 162–4, 226, 228–9
- Robertson, D.H., 127, 145
- Robinson, J., 108–9, 122
- Rochon, L.-P., 126–7
- Roelandts, M., 63
- Rogoff, K.S., 70
- Romer, D., 117–18
- Rowthorn, R.E., 127
- Rueff, J., 207–8, 210
- Samuelson, P.A., 109
- Sargent, T.J., 95, 97
- Sawyer, M., 125
- Say, J.B., 8–9, 38–42, 44, 48, 55, 104, 126, 136–7, 139
- Schmitt, B., 2, 4–8, 11, 24, 26–7, 33, 38, 43–9, 55, 64, 71, 77, 101, 116–17, 121, 138–9, 142, 147, 151, 158, 165–6, 168, 170, 173, 175–6, 179–82, 184–5, 190, 194–6, 201, 204–5, 208, 210–14, 217, 219–24, 230, 233–6, 239, 241, 244, 246–50, 255–8
- Schumacher, E.F., 5, 241, 251–3
- Schumpeter, J.A., 59, 64–5, 67, 126, 130, 152, 157, 162–3, 186, 227
- Schwartz, A.J., 84
- Seager, H.R., 190
- Seccareccia, M., 30–1, 33
- Smets, F., 103, 120
- Smith, A., 2, 21, 29–30, 42, 90, 121, 151–6, 209
- Snowdon, B., 109, 111
- Sowell, T., 41
- Stamp, M., 241
- Stiglitz, J., 110, 117–18
- Stockhammer, E., 127
- Tobin, J., 21, 124
- Turgot, A.R.J., 152–3, 155
- Von Mises, L., 21, 66
- Walras, L., 2, 4, 8, 10, 38, 54, 75, 87–8, 90–1, 98, 100, 104, 117, 122, 141–2, 151, 161–2, 167, 188
- White, W., 79, 241
- Wicksell, K., 66, 199–201
- Woodford, M., 77, 96–8, 102–4, 119
- Wouters, R., 103, 120
- Wray, L.R., 78, 80–1, 124–5, 128
- Young, W., 109, 141

Subject Index

- absolute exchange, 26–7, 40, 43, 121, 131
- aggregate demand
 - and consumers' consumption ability, 135
 - and investors' investment ability, 135
 - and monetary/real shocks, 134
- aggregate supply, 134–5
- alienation, 175, 180

- Bank for International Settlements, 79
 - Bank for International Settlements Triennial Central Bank Survey*, 210
- bank money, 1–2, 3
 - and capital, 197
 - and demand-side economics, 142
 - as dematerialized means of payment, 209
 - and double-entry bookkeeping, 31, 83, 112
 - and gold reserves, 90
 - immaterial nature of, 45, 83
 - importance of, 2
 - international, 251
 - international settlement system use of, 251
 - misconception of, 205–10
 - and monetary system, 138
 - vehicular nature of, 207
- Bank of England, 226
- banks and banking
 - bank credit, 228
 - empty money issued by, 226
 - financial and monetary intermediation of, 227
 - important role of, 165–6
 - as intermediaries, 165, 169, 199
 - separation of income and capital in, 232–8
 - separation of money and credit in, 227–32
- Böhm-Bawerk, E., 10, 66, 78, 151, 158, 160–4, 167, 185–90, 198
 - on capital accumulation, 161–7
 - capitalization approach, 185
 - on circulating capital, 163
 - on fixed capital, 163
 - Leihzins*, 187–9
 - roundabout methods of production, 186–7
 - time preference theory, 186
 - Urzins*, 187–9
- boom-and-bust cycles
 - business cycles vs., 59–82
 - as endogenous phenomena, 78–82
- Brahmagupta, 19
- Bretton Woods conference, 205, 251
- business cycles
 - boom-and-bust cycles vs., 59–82
 - credit cycles vs. real, 76–8
 - derived from statistical observation, 67–70
 - theories of, 64–7
- business-cycle theories, 64–7

- capital, 10
 - circulating, 154–5, 163
 - constant, 159, 163
 - defined, 156, 185
 - financial component of, 166
 - fixed, *see* fixed capital
 - formation as structural cause of economic disorder, 167–75
 - organic composition of, 159
 - pathological, 10
 - and savings, 171, 197
 - and time, 171
 - variable, 159, 163
 - working, 73
- capital accumulation
 - capital formation and economic disorder, 167–75
 - defined, 131
 - and economic crises, 151–83

- capital accumulation – *continued*
 - and inflation, 176–83
 - and unemployment, 176–83
- capital goods, 11, 79, 108
 - circulating, 155
 - and creation of fund, 153
 - fixed, 154–6, 158, 163, 170–5, 234–7
 - production of, 153–4
- capitalism, 7–8
 - and alienation of labour, 175
 - described, 234
 - and labour-power, 160
 - Marx on, 62–3
 - pathological working of, 131
- capitalization, 185–91
- capital loans, 162
 - see also* loans
- capital-time, 169, 196–8
- central bank money, 125
 - see also* State money
- central banks, 140, 242–3
 - see also* banks and banking
- circulating capital, 154–6, 159, 163
- commercial real-estate loans, 81
 - see also* loans
- commodity market, 40, 52–3, 87, 93
- constant capital, 158–9
- consumers' behaviour, 140, 142–3
- consumption
 - goods, demand for, 137, 144
 - variation and total demand, 144
- consumption goods, 42–3, 132
 - as circulating capital, 155–6, 158–9, 163
 - and effective demand, 132, 144
 - production of, 153–4
 - stock of, 235–6
- contract loans, 187, 189
 - see also* loans
- credit
 - and money, 30–7
 - money creation and, 31–3
 - opening of a line of, 20, 24
 - ordinary, 33, 168
 - quantum, 168
 - separation of, and money, 227–32
- credit cycles vs. real business cycles, 76–8
- credit lines, 20, 24, 239, 240
- credit money, 126
 - see also* money
- crises
 - and capital accumulation, 151–83
 - as endogenous disturbances, 60–4
 - and global demand, 130–6
 - and global supply, 130–6
 - and interest, 185–99
 - and interest rates, 199–203
 - from a micro to a macroeconomic analysis of, 70–6
 - and Say's law, 136–40
 - supply-side vs. demand-side economics, 140–8
 - and unemployment, 133
- current demand, 144
- current supply, 144
- debts
 - external, *see* external debts
 - ordinary, 224
 - pathological, 224
- debt servicing, 210–13
- deflation, 41, 97, 133–6, 182–3
- demand
 - aggregate, *see* aggregate demand
 - effective, *see* effective demand
 - global, *see* global demand
 - total and consumption variation, 144
 - virtual, 134
- demand-side economics
 - and consumers' decisions, 140, 142–3
 - and effective demand, 140
 - and money, 146
 - vs. supply-side economics, 140–8
- domestic payment systems, 226–39
 - averting systemic financial crises, 238–9
 - separation of income and capital in banks' books, 232–8
 - separation of money and credit in banks' books, 227–32
- double-entry bookkeeping, 15, 22, 46, 144–5, 220
 - discovery of, 1, 19
 - Walras's conception of, 91

- dynamic stochastic general equilibrium (DSGE) models, 119–20, 145–6
- econometrically validated structural models, 119
- economic crises, *see* crises
- economic models
 aim of, 4
 Debreu's, 98
 Walras's, 98
- economic recovery, 73
- economics
 behavioural, 104
 capital as central concept of, 10
 concept of alienation in, 175
 demand-side vs. supply-side, 140–8
 Keynesian, 107–9
 macroeconomic approach to, 23–4
 new classical, 9, 83, 95–6
 new Keynesian, 117–22
 post-Keynesian, 122–8
 quantum macroeconomics and, 2, 4, 7
 as a science, 2, 75
 structural, 104
 supply-side vs. demand-side, 140–8
- economic systems
 monetary macroeconomics of
 modern, 15–37
 pathological working of, 8, 191
- effective demand
 and demand-side economics, 146
 and full employment, 133
 principle of, 132, 134
- employment, 70, 117, 132–7
 and effective demand principle, 132
 full, 132–3, 135–6
 and productive investment, 185
 and Say's law, 136–40
- empty money, 138, 140, 173–4
- euro-market, 209
- European Central Bank (ECB), 243, 254–6
- exchange-value, 162, 164, 195
- expenditure, 23–4, 44, 51–2
 of income, 169, 171, 177
 of profit, 172, 174, 233
- exports
 net commercial, 208
 real, 224
- external debts
 duplication of indebted countries, 220–2
 pathological duplication, 222
 positive, 216
 servicing, double charge of, 210–13
see also debts
- Fetter, F. A., 185–90, 198
 capitalization approach, 185
 interest analysis, 187–9
 time preference approach, 186
- fiat money, 21
see also money
- financial bubble, 210
- financial circulation deposits, 78–9
- financial crises
 and external debt servicing charge, 210–13
 and interest rates, 199–203
 and international payments' system, 205–10
 and sovereign debt problem, 214–25
 systemic, 238–9
- financial instability hypothesis, 79, 128
- Fisher, I., 81, 84–6, 94, 189–90
 on 'debt-deflation' spiral, 81
 on interest, 189
 on money rate of interest, 189
 quantity equation of, 84, 86
 Seager on interest analysis, 190
 transaction equation of, 84–5
- fixed capital, 52–3, 61, 116
 as accumulated labour, 158
 amortization of, 179–80
 and capital-time, 169–73, 196–7
 vs. circulating capital, 155
 defined, 154
 and financial capital, 191
 formation, 172–3
 production of, 155–6
 and productivity, 177, 200
 and profit investment, 152
see also capital accumulation

- foreign loans, 211, 214
 - see also* loans
- foreign surplus imports, 218–20
- frictional unemployment, 137
- Friedman, M., 66, 76, 83–7, 96
 - monetarism and, 84–7
 - quantity theory of money and, 76, 83, 96
 - on the role played by money, 66, 76
- The General Theory of Employment, Interest, and Money* (Keynes), 39, 41, 123, 132, 134, 136, 145, 166, 191
- global demand
 - and economic crises, 130–6
 - identity between, and global supply, 138
 - Keynes on, 132–3
 - Marx on, 130–1
 - and money income generation, 136
- global supply
 - defined, 147
 - and economic crises, 130–6
 - identity between, and global demand, 138
- gold exchange standard, 205
- goods
 - amortization, 175–6, 179, 181
 - capital, *see* capital goods
 - consumption, *see* consumption goods
 - homogeneous, 18, 75
 - instrumental, 53, 132, 172–3
 - investment, 42–3
 - physically heterogeneous, 32
 - surplus-value, 62–4
- gratuitous money, 138, 140
- Great Depression, 72
- Hayek, F.A., 35, 66–9, 76–8, 87
 - on business cycle fluctuations, 76
 - on centrality of the statistical methods, 67
 - on money, 35
 - on the role played by monetary shocks, 66
- Hicks, J.R., 65–6, 106, 109–13, 116, 122
 - IS-LM model of, 109–17
- hoarded savings, 127
- hoarding, 54
- homogeneous goods, 18, 75
- imports
 - foreign surplus, 218–20
 - net commercial, 207–8, 214, 218–19
 - real, 224
- income
 - logical and factual distinction between money and, 22–30
 - multiplier, 34
- industrial circulation deposits, 78–9
- inflation, 97, 127
 - and capital accumulation, 176–83
 - described, 183
 - and fixed capital formation, 152
 - reducing, 183
 - and unemployment, 176–83
- instrumental goods, 53, 132
- interest
 - capitalization vs. physical productivity, 185–91
 - and economic crises, 185–99
 - Keynes's theory of interest, 191–3
 - quantum macroeconomic analysis of, 194–9
- interest rates
 - defined, 192
 - and economic crises, 199–203
 - and financial crises, 199–203
 - Keynes on, 192–3
- International Clearing Bank (ICB), 252
- International Monetary Fund Currency Composition of Official Foreign Exchange Reserves*, 210
- international monetary system, 241–59
 - functions of the sovereign Bureau, 249–50
 - multinational reforms, 251–7
 - operational functioning of the sovereign Bureau, 250–1
 - single country reform, 244–9
 - world reforms, 251–3, 257–9

- international payment system, 205–10
 bank money, 205–7
 gold exchange standard, 205
 non-system of, 206
- intertemporal general-equilibrium
 foundations, 119
- intrinsic value of money, 30
- investment goods, 42–3
- investments
 forced, 54
 Keynes's identity between saving
 and, 48–55
 of labour, 163, 171
 portfolio, 214
 of profit, 173–4, 178, 180–2, 196–7,
 236
 and savings, 144
- involuntary unemployment, 7, 127,
 180–1, 183
see also pathological unemployment
- IS-LM model, 109–17
- Juglar, C., 59–60, 64, 67, 76
 business-cycle theories and, 64–7
 “law of crises”, 67
 on monetary and commercial crises,
 59–60
- Keynes, J.M., 16–19, 22–4, 39, 41–3,
 48, 50–3, 55, 78–9, 106–28,
 132–4, 136, 139, 145, 147–8, 151,
 153, 161, 165–7, 191–3, 198, 226,
 232, 242, 251–5, 258
 analysis of capital, 151
 on capital accumulation, 161–7
 as the ‘father’ of modern
 macroeconomics, 106, 111
 identity between saving and
 investment, 48–55
 identity between Y and $C+I$, 41–4
 on labour as sole factor of
 production, 24
 on nominal money, 18–19
 principle of effective demand, 132,
 134
 on saving and investment
 relationship, 184–5
 on Say's law, 41, 136
 theory of interest, 191–4
 on using international bank money
 for trade deficits, 251–2
 on value of money, 79
- Keynesian economics, 107–9
 post-, 122–8
- labour
 alienation of, 175
 determination of value by, 156
 investment in, 157–8
 Keynes on, 24
 vs. labour-power, 130–1
 Marx on, 61–4
 time, 160
- labour-power, 61–4, 130–1, 158–60,
 175
- labour theory of value, 159–60
- Lehman Brothers, 128
- Leihzins*, 187–8
- loanable funds, 191
- loans
 capital, 162
 commercial real estate, 81
 contract, 187, 189, 199
 contract-interest on, 184
 foreign, 211, 214
 interest on, 187–9
- Lucas, R.E., Jr., 66–8, 71, 76, 95–8, 100,
 118–19
 on microeconomic foundations of
 macroeconomics, 71
 rational expectations hypothesis of,
 96, 118
- macroeconomic laws, 38–55
 identity between each agent's sales
 and purchases, 44–8
 Keynes's identity between saving
 and investment, 48–55
 Keynes's identity between Y and
 $C+I$, 41–4
 of monetary production economies,
 38–55
 Say's law, 39–41
 Schmitt's law, 8, 38, 45–9, 139
- macroeconomic payments, 218
- macroeconomic price, 71
see also prices

- macroeconomics
 - monetary, 15–37
 - quantum, 71, 100, 109
 - quantum monetary, 5
- macroeconomic saving, 10, 52–4, 116, 144–5, 184–5, 196–9
- Marx, K., 10, 23, 29, 60–3, 130–1, 151, 156, 158–60, 163, 167, 175, 201
 - on capital accumulation, 10, 156–61
 - on distinction between nominal and real money, 29
 - on exchange-value, 130
 - falling rate of profit and overproduction, concepts of, 60–4
 - realization problem, 130
 - surplus-value, 130–1
- microeconomic payments, 218
- microeconomic prices, 71
 - see also* prices
- microeconomic saving, 197
- modern economic systems, 15–37
- monetarism, 84–95
 - argument against, 88
 - concepts proper to, 89
- monetary homogeneity, 5
- monetary macroeconomics, 15–37
 - about money, 16–18
 - of modern economic systems, 15–37
 - money and credit, 30–7
 - money and income, logical and factual distinction between, 22–30
 - nominal money, nature of, 18–20
 - value of money, 20–2
- monetary payment, 212–13
- monetary production economies, 38–55
- monetary reforms
 - multinational reforms, 251–7
 - single country reforms, 244–9
 - world reforms, 251–3, 257–9
- monetary unification, 216
- money
 - credit, 126
 - and credit, 30–7
 - defined, 2, 18
 - and demand-side economics, 146
 - described, 16–18
 - empty, 138, 140, 173–4
 - fiat, 21
 - flow nature of, 29, 115
 - gratuitous, 138
 - issued by banks, 206
 - logical and factual distinction
 - between income and, 22–30
 - nature of nominal, 18–20
 - net-asset definition of, 5
 - as a numerical unit, 18
 - as object of exchange, 209
 - purchasing power of, 22, 27–8, 30; *see also* purchasing power
 - separation of, and credit, 227–32
 - State, 20–1, 125
 - value of, 20–2
- money multiplier, 33–4
- money neutrality, concept of, 101
- money or market rate of interest, 200
- moveable wealth, 155
- multinational reforms, 251–7
- natural rate of interest, 199–200
- new classical analysis
 - and rational expectations hypothesis, 95–8
- new classical synthesis, 83–105
 - and rational expectations hypothesis, 95–8
 - from real business cycle analysis to, 98–105
- New Keynesian economics, 117–22
 - and demand-side vs. supply-side economics, 145
 - real business cycle (RBC) models, 145–6
- nominal money
 - nature of, 18–20
 - and real money, 258
 - Smith on, 29
 - and wages, 173
 - see also* money
- Observations on a Paper by Saint-Pérayve* (Turgot), 152
- On the Principles of Political Economy and Taxation* (Ricardo), 151, 156
- ordinary debt, 224
- output, *see* production

- parsimony, 153
- pathological capital, 10
see also capital
- pathological debt, 224
- pathological unemployment, 137
see also involuntary unemployment
- physically heterogeneous goods, 32
- Plan for the Establishment of a National Bank* (Ricardo), 226
- positive interest, 198
- positive purchasing power, 22, 27, 30–1, 35, 234
see also purchasing power
- post-Keynesian economics, 122–8
- prices
 backward indexation of, 120
 macroeconomic, 71
 microeconomic, 71
- production, 44, 121–2
 as absolute exchange, 131
 of amortization goods, 175–6, 179, 181
 Böhm-Bawerk methods of, 167
 defined, 137, 146, 168
 and demand, 129
 of instrumental goods, 172–3
 macroeconomic cost of, 147
 monetary economy of, 152
 and supply, 129
 and supply-side economics, 140, 141
- profit
 expenditure of, 172, 174, 233
 and fixed capital, 152
 investment, 173–4, 178, 180–2, 196–7, 236
 maximization, 140
- prosperity, 65, 73
- purchase
 of a claim, 46
 of deposit certificates, 46
- purchasing power
 of money, 22, 27–8, 30
 positive, 22, 27, 30–1, 35, 234
- quantity theory of money, 54, 83–6, 94–6
- quantum credit, 168
see also credit
- quantum macroeconomics, 71, 100, 109
 analysis of interest, 194–9
 and global supply and demand, 146–7
 and supply-side and demand-side economics, 147–8
 theoretical frameworks of, 129
- quantum monetary macroeconomics, 5
- rate of interest, *see* interest rates
- rational expectations hypothesis
 and new classical analysis, 95–8
- real business cycle (RBC) models, 145–6
- real business cycles
 vs. credit cycles, 76–8
 new classical synthesis and, 98–105
- realization problem, 130
- real payment, 212–13
- recession, 73, 127
- relative exchanges, 27, 39, 45, 89, 92, 104–5, 121
- Report of the UN Commission of Experts*, 242
- Ricardo, D., 10–11, 29, 151, 156–9, 162–4, 226, 228–9
 on capital accumulation, 10, 156–61
 on determination of value, 156
 on distinction between nominal and real money, 29
 and Marx, 156–61
Plan for the Establishment of a National Bank, 226
 separation of money and credit in banks' books, 227–32
- saving(s), 167
 accumulated, 33
 and capital, 171, 197
 concept of, 167
 hoarded, 127
 Keynes's identity between investment and, 48–55
 macroeconomic, 10, 52–4, 116, 144–5, 184–5, 196–9
 microeconomic, 197

- Say's law, 39–41
 and economic crises, 136–40
 implications on employment, 136–40
 Keynes on, 136
- Schmitt, B., 2, 4–8, 11, 24, 26–7, 33, 38, 43–9, 55, 64, 71, 77, 101, 116–17, 121, 138–9, 142, 147, 151, 158, 165–6, 168, 170, 173, 175–6, 179–82, 184–5, 190, 194–6, 201, 204–5, 208, 210–14, 217, 219–24, 230, 233–6, 239, 241, 244, 246–50, 255–8
- absolute exchange, concept of, 26–7, 43
- on bank money, 2
- on capital accumulation, 179
- capital-time, 196
- on external debts, 214–15
- on interest, 194–5
- macroeconomic analysis of, 4
- on monetary reforms, 234, 239
- pathological capital, 180
- on payment of wages, 139
- quantum analysis of, 44, 72
- on quantum vs. ordinary credit, 168
- single country reform, 244, 246–7, 256
- and sovereign debt crisis, 222–4
- world monetary reform, 258
- Schmitt's law, 8, 38, 45–9, 139
- Schumpeter, J.A., 59, 64–5, 67, 126, 130, 152, 157, 162–3, 186, 227
- on Böhm-Bawerk's physical capital, 163
- on Turgot's role in capital theory, 152
- Seager, H.R.
- on capital, 190
- on Fisher's interest analysis, 190
- single country reforms, 244–9
- Smith, A., 2, 21, 29–30, 42, 90, 121, 151–6, 209
- on capital accumulation, 152–6
- on circulating capital, 154–5
- on fixed capital, 154–5
- money definition of, 29
- on money's purchasing power, 30
- nominal and real money, distinction between, 2, 29
- sovereign Bureau
- functions of, 249–50
- operational functioning of, 250–1
- sovereign debt, 214–25
- defined, 222–3
- foreign surplus imports, 218–20
- indebted countries' external debts
- duplication, 220–2
- pathological nature of, 222–5
- Special Drawing Rights (SDRs), 242–3
- State money, 20–1, 125
- see also* central bank money; money
- structural unemployment, 125
- supply
- aggregate, 134–5
- global, *see* global supply
- virtual, 134
- supply-side economics
- vs. demand-side economics, 140–8
- and productivity, 140, 141, 146–7
- and profit maximization, 140
- surplus-value, concept of, 61–4, 130–1
- surplus-value goods, 62–4
- systemic financial crises, 238–9
- Tableau économique*, 152
- tautology, defined, 85
- theorem of interest, 204–5
- Traité d'économie politique* (Say), 39
- A Treatise on Money* (Keynes), 16, 123, 145, 165, 166, 226, 234
- Turgot, A.R.J.
- on capital accumulation, 152–6
- unemployment, 61
- and capital accumulation, 176–83
- causes of, 134–5
- concepts of, 133
- defined, 133
- and demand, 132
- and economic crises, 133
- frictional, 137
- and inflation, 176–83
- involuntary, 7, 127, 180–1, 183
- pathological, 137
- structural, 125
- Urzins*, 187–8

- value, determination of, 156
- value of money, 20–2
 - intrinsic, 30
- variable capital, 159
- virtual demand, 134
- virtual supply, 134
- wages
 - Marx on, 23
 - payment of, 25–35
 - result of the payment of, 26
 - surplus-value and, 61
- Walras, L., 2, 4, 8, 10, 38, 54, 75, 87–8,
 - 90–1, 98, 100, 104, 117, 122,
 - 141–2, 151, 161–2, 167, 188
 - on capital accumulation, 10, 161–7
 - double-entry bookkeeping,
 - conception of, 91
 - on economic value, 2
 - Elements of Pure Economics*, 91
 - general equilibrium model of, 4, 88,
 - 100
 - neoclassical general equilibrium
 - approach, 188
 - Walras's law, 141–2
 - The Wealth of Nations* (Smith), 30, 154
 - Wicksell, K.
 - analysis of interest and prices,
 - 199–200
 - natural rate of interest, definition
 - of, 199–200
 - working capital, 73
 - see also* capital
 - world reforms, 251–3, 257–9
- Wray, L.R., 78, 80–1, 124–5, 128
 - 'modern money theory' of, 80
 - on purchasing power of money, 124
- zero-sum transaction, 168–9

