CONTEMPORARY TOPICS IN MACROECONOMICS

Edited by
Julio César Leal Ordóñez & Stephen McKnight

339.091724
C7617

EL COLEGIO DE MÉXICO

JULIO CÉSAR LEAL ORDÓÑEZ is currently a research economist at Banco de México. He received his Ph.D. in economics from Arizona State University, and has previously worked as a *profesor-investigador* at CIDE. In 2011, Julio was awarded the Banamex Prize in Economic Research.

STEPHEN MCKNIGHT is currently a *profesor-investigador* at El Colegio de México. He received his Ph.D. in economics from the University of Essex, and has previously worked as an assistant professor at the University of Reading. Recent publications have appeared in *International Finance* and *Oxford Economic Papers*.

Contemporary Topics in Macroeconomics

CENTRO DE ESTUDIOS ECONÓMICOS

Contemporary Topics in Macroeconomics

Edited by
JULIO CÉSAR LEAL ORDÓÑEZ
and
STEPHEN MCKNIGHT

EL COLEGIO DE MÉXICO

339.091724
C7617
Contemporary topics in macroeconomics / edited by Julio César Leal Ordóñez and Stephen McKnight. — 1ª ed. — México, D.F. : El Colegio de México, Centro de Estudios Económicos, 2014.
183 p. ; 22 cm.

ISBN 978-607-462-637-7

1. Macroeconomía — Países emergentes — Modelos econométricos. 2. Equilibrio (Economía) — Países emergentes — Modelos econométricos. 3. Países emergentes — Condiciones económicas — Siglo XXI — Modelos econométricos. I. Leal, Julio, ed. II. McKnight, Stephen, ed.

DDC-23

Los puntos de vista expresados por los autores de los capítulos que conforman este libro son de responsabilidad exclusiva de sus autores y no representan la posición institucional de El Colegio de México.

Primera edición, 2014

D.R. © El Colegio de México, A.C.
Camino al Ajusco 20
Pedregal de Santa Teresa
10740 México D. F.
www.colmex.mx

ISBN 978-607-462-637-7

Impreso en México

CONTENTS

FOREWORD

Contemporary Topics in Macroeconomics, edited by Julio César Leal Ordóñez and Stephen McKnight, constitutes an important intellectual effort in adapting macroeconomic models of general equilibrium to the problems of developing countries. The publication of this book is a source of deep satisfaction for the Centre for Economic Studies, as it is the result of the collaboration between a number of prestigious Mexican institutions: Banco de México, Centro de Investigación y Docencia Económicas, El Colegio de México, and Instituto Tecnológico Autónomo de México.

This year, the Centre for Economic Studies celebrates its fiftieth anniversary. During this time, the Centre has been driven by the desire to provide international quality training in economics to talented students who, it is envisaged, will improve the future conditions of our country. This is supported by the day-to-day research activities undertaken by the professors and associates of the Centre. The present-day work of the Centre strives to uphold the same standards established five decades ago.

Of course, the aims and objectives of the Centre could not be achieved without the generous help and support of a number of institutions and partners, including Banco de México. Indeed, Banco de México was a founding member of El Colegio de México in 1940, and ever since, it has worked with us both in the training of students and the promotion of economic research. In addition, Banco de México is a member of the committee for the Víctor L. Urquidi annual prize in economics.

I hope that this collaboration is the first of many inter-institutional projects in the future. I congratulate the editors and the authors who participated for their accomplishment and fully expect the book to become an invaluable teaching and research resource.

José Romero
Director
Centre for Economic Studies
El Colegio de México

PREFACE

This book presents a selection of papers on modern macroeconomics. It originated from a series of discussions with a number of colleagues relating to the areas of macroeconomic research currently being undertaken by academics in Mexico. In early 2013 we invited a group of research-active macroeconomists to discuss their current topics of interest. The present volume is the result.

The reflexive criticism of macroeconomics that unfolded after the financial crisis of 2008, especially the attacks directed against the use of dynamic stochastic general equilibrium models, was a crucial motivating factor for writing this book. This was augmented by the neglect of many economic commentators to the several macroeconomic crises that had occurred in developing countries during the 1980's and 1990's. Consequently, we saw the need for a collective effort to offer a balanced view of macroeconomics. We wanted to present a selection of current macroeconomic research based on dynamic stochastic general equilibrium models, where the issues discussed were not only relevant for developing countries, but informative to the discussions currently being held in developed countries. Of course, the reader will be the judge of how successful we have been.

The book comprises of six self-contained chapters written by the invited contributors covering a number of issues in macroeconomics, all using the techniques of dynamic general equilibrium. The topics covered in this collection range from the impact of the informal sector for the macroeconomy; the aggregate effects of changes in taxation and social expenditure; an examination of business cycles in both developed and developing countries; an investigation of liquidity traps under financial instability; and the design of monetary policy from the perspective of small open economies. We hope that the range and diversity of topics addressed in the chapters that follow will appeal to a broad audience, including students, policymakers, and academic researchers alike.

Julio César Leal Ordóñez
and *Stephen McKnight*

México, D.F., 2014.

ACKNOWLEDGEMENTS

We would like to thank Dr. Alejandro Castañeda Sabido who helped concoct the initial idea, and Dr. José Romero for providing the encouragement and support so that it actually happened. We also thank Roberto Gómez and Erick Martínez for excellent research assistance, and the typesetter Axel E. Retif for a superb job. We are especially grateful to Dr. Joàna Chapa Cantú and Dr. Alexander Mihailov for providing very helpful comments and feedback on an earlier draft of the manuscript. Leal also thanks the members of the Dirección General de Investigación Económica at Banco de México for their support during the process, especially Dr. Ana María Aguilar. Finally, we are indebted to the authors that contributed to the project: the book is the result of their efforts.

LIST OF CONTRIBUTORS

Jorge Alonso-Ortiz (jorge.alonso@itam.mx) is currently an assistant professor in the Centro de Investigación Económica at the Instituto Tecnológico Autónomo de México (ITAM). He received his Ph.D. in economics from Arizona State University in 2010 under the supervision of Richard Rogerson. His research interests include macroeconomics, labor economics, and applied econometrics. Jorge has a recent publication in the *Journal of Monetary Economics.*

Arturo Antón Sarabia (arturo.anton@cide.edu) is currently a *profesor-investigador* at the Centro de Investigación y Docencia Económicas (CIDE). He received his Ph.D. in economics from Cornell University and previously worked as an economist at Banco de México. His research interests include dynamic general equilibrium, business cycles, and monetary and fiscal policy. He is a coauthor of the book *The End of Informality in Mexico? Fiscal Reform for Universal Social Insurance.* Recent publications include *Economía Mexicana, Energy Policy,* and *El Trimestre Económico.* Arturo is a member of the Sistema Nacional de Investigadores.

Julio A. Carrillo (jcarrillo@banxico.org.mx) is a research economist at Banco de México. He received his Ph.D. in economics from the Toulouse School of Economics, and has previously worked as an assistant professor at Maastricht University, in the Netherlands, and as a post-doctoral researcher at Ghent University, in Belgium. His research interests include macroeconomic theory, monetary economics, and applied macroeconometrics. Recent publications have appeared in the *Journal of Economic Dynamics and Control* and the *Review of Economic Dynamics.*

Julio César Leal Ordóñez (jleal@banxico.org.mx) is a research economist at Banco de México. He received his Ph.D. in economics from Arizona State University in 2010, and has previously worked as a *profesor-investigador* at CIDE. His research interests include macroeconomics, labor economics, and development economics. In 2011, Julio was awarded the Banamex Prize in Economic Research. Recent publications have appeared in the *Review of Economic Dynamics.* Julio is a member of the Sistema Nacional de Investigadores.

Stephen McKnight (mcknight@colmex.mx) is currently a *profesor-investigador* at El Colegio de México. He received his Ph.D. in economics from the University of Essex, and has previously worked as an assistant professor at the University of Reading, in the United Kingdom. His research interests include open-economy macroeconomics, business cycle theory, and monetary economics. Recent publications have appeared in *International Finance* and *Oxford Economic Papers*. Stephen is a member of the Sistema Nacional de Investigadores.

Céline Poilly (celine.poilly@unil.ch) is currently an assistant professor at the University of Lausanne, in Switzerland. She received her Ph.D. in economics from the University of Cergy-Pontoise and the Banque de France, and has previously worked as an assistant professor at the Université catholique de Louvain, in Belgium. Her research interests include macroeconomic theory, monetary and financial economics, and macroeconometrics. Recent publications have appeared in the *Journal of Monetary Economics*, *Macroeconomic Dynamics*, and the *Review of Economic Dynamics*.

I. REFLECTIONS ON MODERN MACROECONOMICS

JULIO CÉSAR LEAL ORDÓÑEZ[1]
Banco de México

STEPHEN MCKNIGHT
El Colegio de México

I.1 INTRODUCTION

Modern macroeconomics is a vast, rapidly expanding discipline with many diverse applications. At the core of modern macroeconomic research is the use of dynamic stochastic general equilibrium (DSGE) models. These are intertemporal models based on microeconomic foundations in which rational agents such as households and firms make optimizing decisions.[2] Ever since Kydland and Prescott (1982) used the equilibrium growth model to explain the cyclical properties of aggregate variables, DSGE models have been employed to study a broad set of macroeconomic issues. While these models are often associated with the study of monetary policy, they are currently being used to investigate all kinds of exciting and interesting economic questions that go well beyond the field of monetary economics. This book is a sample of some of the macroeconomic research currently being undertaken using the techniques of dynamic general equilibrium.

While the themes and methods chosen are appropriate for both developed and developing economies, one particularity of the book is its emphasis on the economic problems and challenges faced by policymakers in the developing world. As is argued in the sections below, it is our belief that more attention in the literature should be placed on the economic issues of developing countries, and its associated research.

Of course, it is impossible to talk about modern macroeconomics without some reference to the recent financial crisis of 2008 and the widespread criticism (both from inside and outside the economics profession) directed at the discipline, and in particular, the dynamic general equilibrium approach. This is where our discussion begins.

[1] The views expressed in this chapter are those of the authors, and do not necessarily reflect those of Banco de México.

[2] For a non-technical discussion see Woodford (2009).

I.2 The financial crisis and the state of macroeconomics

I.2.1 The financial crisis of 2008

Since the 1980's, the world economy had experienced a prolonged period of economic stability. The so-called *great moderation* witnessed the steady decline in both the variability of world inflation (see Figure I.1) and the volatility of world output (see Figure I.2).[3] The financial crisis of 2008 brought this era of macroeconomic stability to an abrupt end. As illustrated in Figure I.2, the crisis represented an unprecedented episode of global output volatility. Figure I.3 depicts annual world GDP growth rates for the period 1980–2012. As Figure I.3 shows, the 2008 crisis triggered a collapse in world growth. The adverse economic effects of this severe global economic downturn spread to labor markets, resulting in substantial increases in worldwide unemployment.

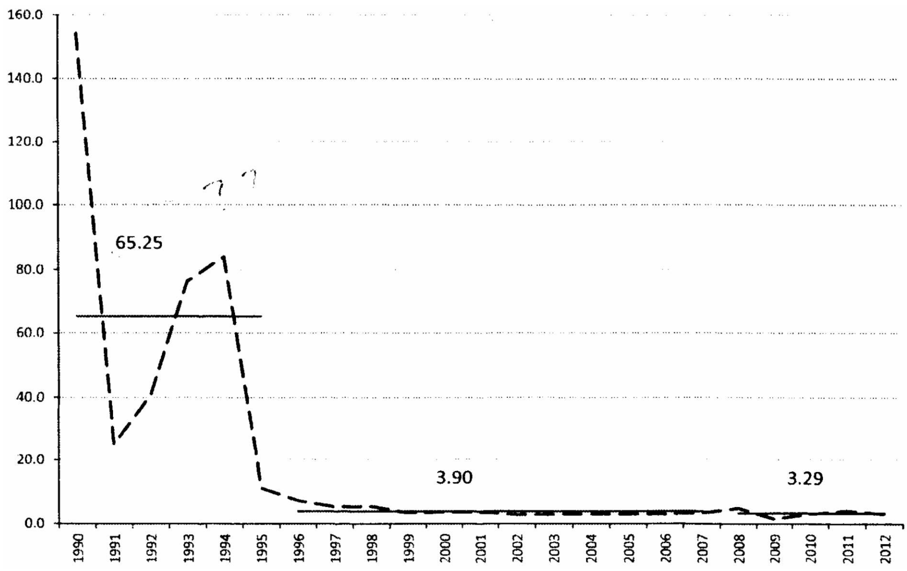


Figure I.1. World inflation (1990–2012)
Source: Banco de México

[3] Output volatility is defined as the absolute difference of the growth rate in each period minus the average growth rate of the interval. The intervals are 1980–1985, 1986–2007, and 2008–2012.

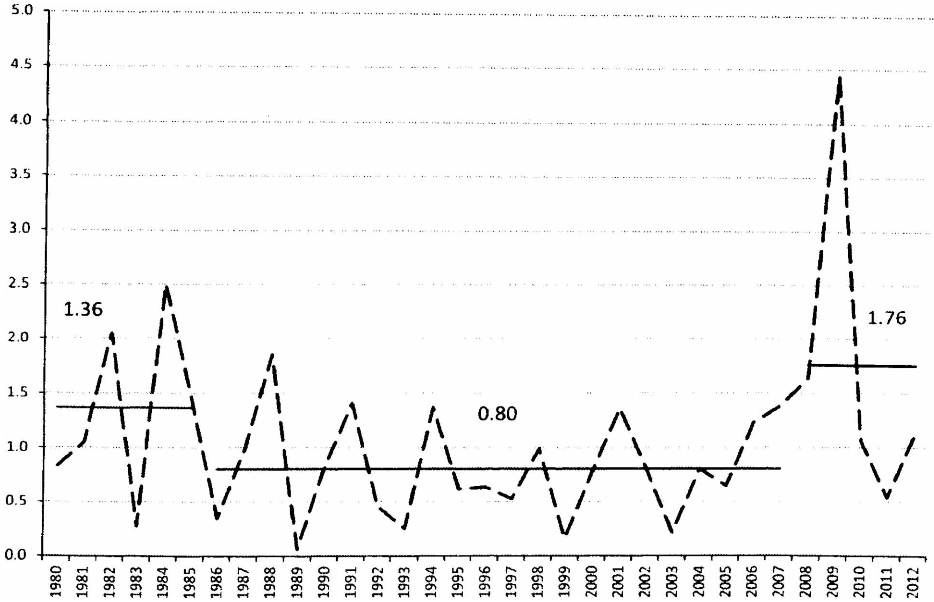


Figure I.2. Volatility of world GDP (1980–2012)
Source: Banco de México with information from INEGI

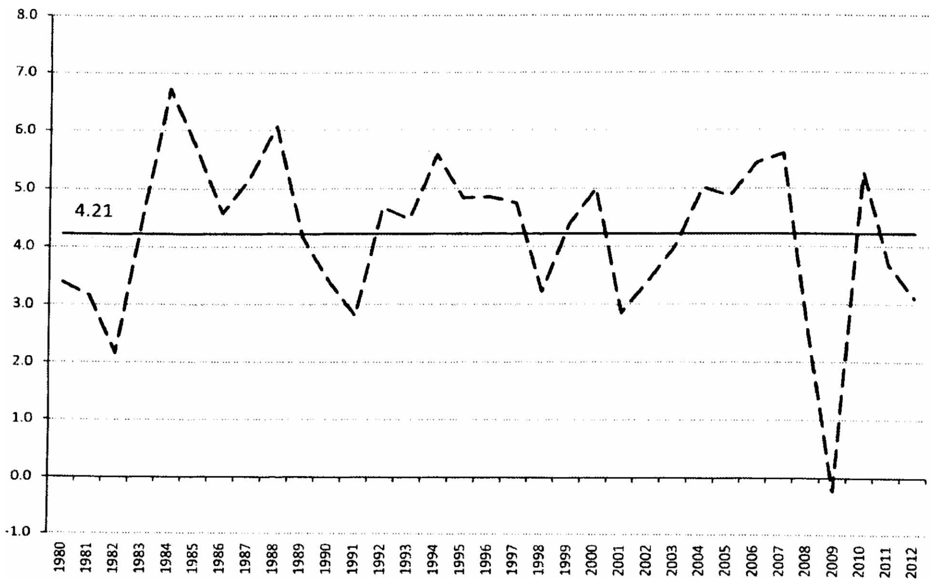


Figure I.3. World GDP growth rates (1980–2012)
Source: Banco de México

1.2.2 A crisis for macroeconomics?

The 2008 financial crisis also resulted in widespread criticism of modern macroeconomics and the methodologies it employs. The criticism came not only from outside the economics profession, but also from several leading macroeconomists. For example, Acemouglu (2009) wrote about "the lessons from our intellectual complaisance", and Caballero (2010) discussed "the pretense of knowledge". However, the most stinging criticisms were directed to the use of dynamic general equilibrium models in modern macroeconomics:

> The central cause of the profession's failure was the desire for an all-encompassing, intellectually elegant approach that also gave economists a chance to show off their mathematical prowess. [Krugman, 2009]

> The so-called dynamic stochastic general equilibrium approach has become so mesmerized with its own internal logic that it has begun to confuse the precision it has achieved in its own world with the precision about the real one. [Caballero, 2010]

> The standard macroeconomic models have failed, by all the most important tests of scientific theory. They did not predict that the financial crisis would happen; and when it did, they understated its effects. ... the sum of these failures points to the need for a fundamental re-examination of the models. [Stiglitz, 2011]

There is little doubt that during the great moderation DSGE models were guilty of ignoring financial market frictions.[4] Prior to the crisis, popular macroeconomic models assumed frictionless perfect financial markets in every aspect of the economy (Rogoff, 2010). Therefore, it is unsurprising that these models did not predict the financial crisis and offered little, if any, guidance for policymakers. This ignorance of financial frictions during the great moderation is also unsurprising since macroeconomists (especially in the developed world) were using DSGE models to explain different research questions:

> The criticism ... is misdirected and reflects a misunderstanding of the purpose for which those models were devised. These models were designed to describe aggregate economic fluctuations during normal times when markets can bring borrowers and lenders together in orderly ways, not during financial crisis and market breakdowns. [Sargent, 2010]

[4] Many DSGE models were incorporating frictions in the labor and product markets via the introduction of price and wage rigidities.

Furthermore, the absence of financial frictions from DSGE models does not imply that the discipline as a whole did not understand the importance of financial stability. Caballero (2010) makes an interesting division of macroeconomic research between *the core* (i.e. DSGE models) and *the periphery* (i.e. models that ignore complex general equilibrium interactions and focus on the details of the sub-problems and mechanisms). Paraphrasing Caballero (2010), the periphery has made important progress in understanding phenomena like speculative bubbles, leverage cycles, liquidity runs, among many other things. As also pointed out by Sargent (2010), there were at least two very influential models about financial stability —Karaken and Wallace (1978) and Diamond and Dybvig (1983)— that stressed the importance of government policy and regulation to prevent bank-runs and moral-hazard problems.

While the recent crisis may have been a wake-up call for some economists, for those interested in the macroeconomics of developing countries such issues are all too familiar. Indeed, there is a huge literature that has attempted to understand episodes of economic and financial crises in developing economies. For instance, the problem of *sudden stops* in the flow of external finance pioneered the study of financial crises within the context of DSGE models.[5] An important lesson for the discipline arising from the 2008 crisis is that research focusing on the problems of developing countries should not be automatically disregarded as being irrelevant for developed countries; rather it is our view that such research can be highly informative.

Overall, in our opinion the claim that macroeconomists have nothing to say about economic crises and financial (in)stability is simply not correct. Yet there is no doubt that there is a need for *the core* of macro to incorporate both financial frictions and an explicit banking sector into monetary DSGE models.[6]

> The recent crisis has reminded us that the financial sector may be both a source of shocks and an amplification mechanism for these shocks, with potentially important consequences for the economy and for the effectiveness of policy tools. [Vergara, 2013]

The discipline has already made some significant advances on this front. To date, the research on financial frictions has taken two main approaches. One approach has integrated financial frictions between *lenders and borrowers* into an otherwise standard DSGE framework, whereas the second approach has embedded *inter-bank lending* frictions. There are two main ways to incorporate frictions

[5] For an excellent review of this literature see Arellano and Mendoza (2002). An important caveat of this literature is that the key features used to capture financial frictions were not micro-founded.

[6] There is also a desire among some researchers to improve the way expectations of the future are modeled in intertemporal general equilibrium models. See, e.g., Woodford (2012).

between borrowers and lenders. One way is via collateral constraints (based on Kiyotaki and Moore, 1997); the other is through an emphasis on monitoring costs and an external finance premia (based on Carlstrom and Fuerst, 1997; and Bernanke et al., 1999). Incorporating a banking sector into a DSGE model that is useful to study aggregate fluctuations is a challenge. The most promising research in this area is perhaps the work done by Gertler and Kiyotaki (2010) that includes a full banking sector, and where small distortions can be amplified via frictions between banks and depositors.

For policymakers, the recent crisis created many challenges. At the microeconomic level, in an attempt to avoid similar moral hazard problems occurring again, and to aid those sectors most affected by the crisis, policymakers have changed the regulatory framework faced by firms. As documented by McGrattan and Prescott (2012), in order to comply with these regulations, the costs paid by business have risen dramatically in recent years. Such increases in regulatory burden could have a significant negative impact on productivity. Important challenges have also arisen for monetary policy.[7] For example, as central banks reduced interest rates in response to falling output, policy rates were quickly driven to near-zero in many developed countries. Consequently, policymakers have experimented with *unconventional* monetary policy tools in an attempt to stimulate the economy. One policy that has received considerable interest is *forward guidance*, whereby the central bank generates stronger incentives for spending today by creating expectations of looser monetary policy in the future than would otherwise have been expected (Woodford, 2013). In Chapter V of this book, Julio A. Carrillo and Céline Poilly develop a DSGE model that incorporates financial frictions as in Bernanke et al. (1999). Among other things, they use their model to examine the effectiveness of forward guidance when interest rates are near zero. They find that while such a policy can have benefits in reducing the severity of the recession, the consequences of miscalculating and providing abundant liquidity for longer than is necessary can have serious destabilizing effects on the economy. This is an important contribution to the literature, which will stimulate further debate on the costs and benefits of adopting a forward-guidance policy.

However, as we stressed earlier, this book is not only about the recent financial crisis. Macroeconomics, and the dynamic general equilibrium approach, is much more diverse. In this collection we aim to also highlight some of the other exciting and interesting research topics that are currently being investigated, many of which are of particular importance for developing countries. This is a topic we now address in the following section.

[7] For an excellent discussion of these issues see Blanchard et al. (2010) and Reichlin and Baldwin (2013).

I.3 The Macroeconomics of Developing Countries

I.3.1 Business cycle fluctuations

The economic problems in developed countries differ from those faced by developing countries in several dimensions. One major difference relates to the nature of business cycles. As discussed in detail by Arturo Antón Sarabia in Chapter II of this book, there are a number of stylized differences between developed and developing economies at business cycle frequencies. For example, it is well-documented that output volatility is much higher on average in the developing world. This is illustrated by Figures I.4 and I.5, which show de-trended real GDP per capita during the period 1980–2012 for two representative groups of developed and developing countries.[8] By inspection of Figures I.4 and I.5, output fluctuates much more in emerging economies than developed economies.

Table I.1. Mean % standard deviation of GDP (1980–2012)

	Developed countries	*Developing countries*
1980–2012	1.47	2.87
1980–2000	1.39	3.12
2001–2012	1.59	2.12

Source: Author's own calculations using IMF-WEO data.

This can also be seen from Table I.1 where we have computed the average percentage standard deviation of the cyclical component of output for the periods 1980–2012, 1980–2000, and 2001–2012. The sample consists of 29 developed economies (20 from Europe, 5 from Asia, 2 from North America, and 2 from Australia and Oceania) and 106 emerging and developing economies (39 from Sub-Saharan Africa, 27 from Latin America and the Caribbean, 16 from North Africa and the Middle East, 13 from Asia, 6 from Australia and Oceania, and 5 from Europe). For the period 1980–2012, the mean standard deviation of GDP for developed countries is 1.47%, whereas for emerging economies it is 2.87%. Consequently, for our sample the mean standard deviation of output is over 1.9 times higher in developing countries. Table I.1 also reports averages of the percentage

[8] We computed the cyclical component by first taking logs and then detrending the series using a HP filter with smoothing parameter equal to 6.25.

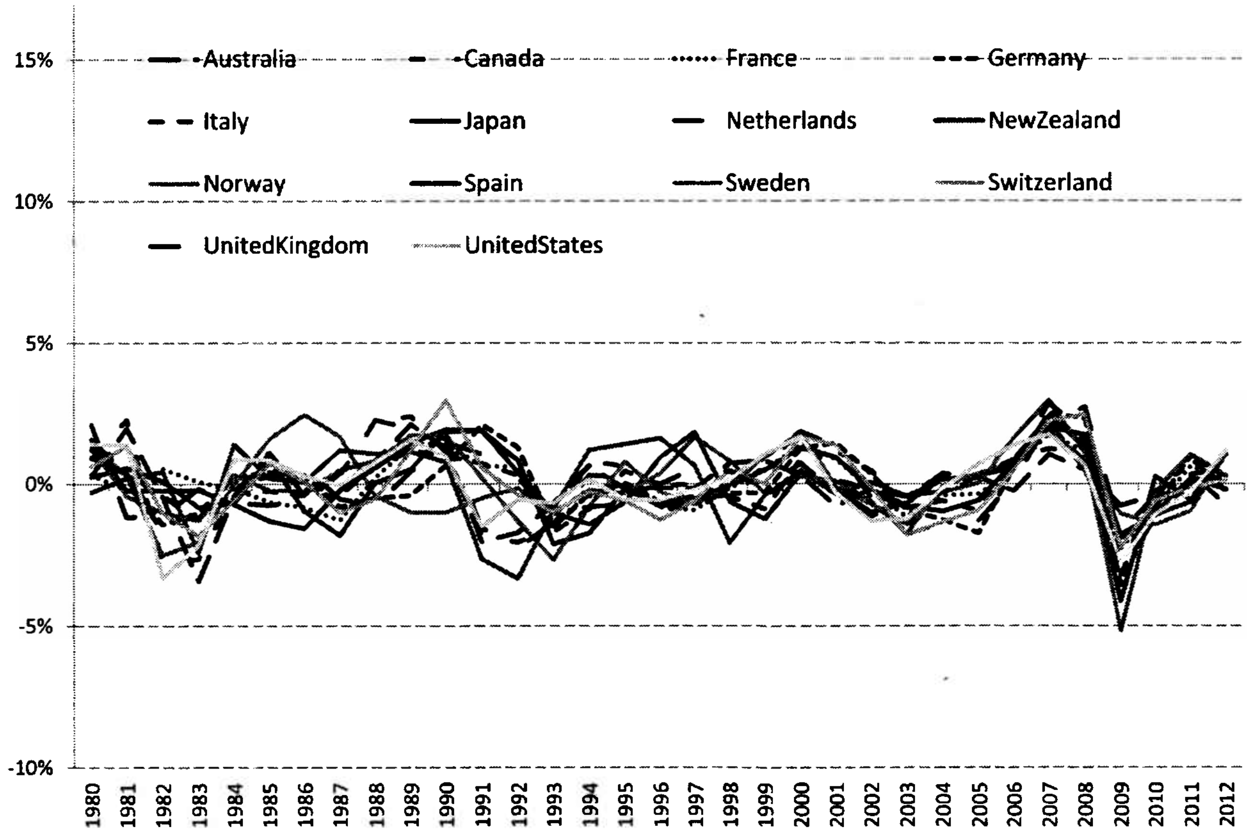


Figure I.4. GDP fluctuations (1980–2012): Selected developed countries
Source: Author's own calculations using IMF-WEO data

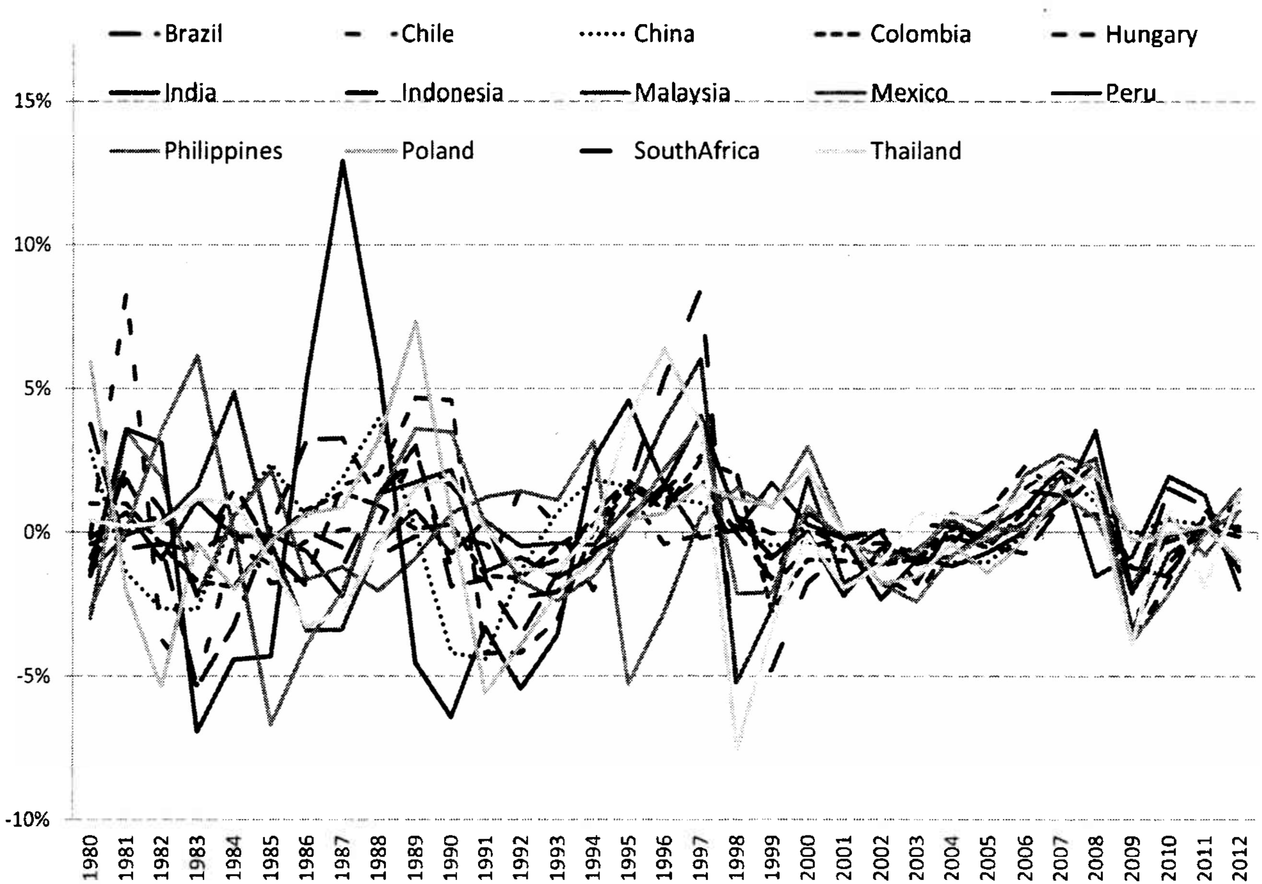


Figure I.5. GDP fluctuations (1980–2012): Selected developing countries

Source: Author's own calculations using IMF-WEO data standard deviation of output for the sub-periods 1980–2000 and 2001–2012. Interestingly these results suggest, at least for the sample studied here, that the volatility of output in developing countries has decreased significantly over the last decade.

The higher volatility of output in developing countries is related to the fact that financial, debt, and current-account crises are typically more frequent in these economies than in developed countries. One of the major challenges of the *macroeconomics of developing countries* has been to construct a coherent model that can explain the main features associated with the various crises experienced by emerging economies. For example, sudden drops in output, combined with current-account reversals and credit-crunch episodes, have been hard to explain using frictionless models. Consequently, this literature has introduced a number of different frictions to help explain business cycles in emerging economies (see, e.g., Neumeyer and Perri, 2005; Aguiar and Gopinath, 2007; García-Cicco et al., 2010; Mendoza and Yue, 2012). As discussed by Arturo Antón Sarabia in Chapter II of this book, one strand of this literature has introduced efficiency and labor frictions into intertemporal models in order to explain particular features of the business cycle in developing countries.

Another strand of the literature has concentrated instead on financial frictions. For example, Mendoza and Yue (2012) focus on episodes of sovereign default. These episodes are an important component of the business cycles in many emerging economies, and are associated with large decreases in Total Factor Productivity (TFP). To illustrate this, Figure I.6 compares measured TFP growth for Mexico and the U.S. during the period 1980–2012.[9] By inspection of Figure I.6, Mexico has experienced large falls in TFP during episodes of debt crises: most notably the 1982 crisis and the 1994 *tequila* crisis. Such large reductions are not typically common in the business cycles of developed countries and have proven to be difficult to explain in the literature.

On the one hand, a branch of the literature has developed models that assume that all shocks experienced by emerging economies can be summarized by *exogenous* TFP shocks (see, e.g., Kydland and Zarazaga, 2002). On the other hand, models have been developed that *endogenize* TFP (see, e.g., Mendoza and Yue, 2012; Kim, 2014). For instance, Mendoza and Yue (2012) argue that there is an amplification mechanism where the presence of financial frictions and imperfect substitutability between imported and domestic inputs play a crucial role. In

[9] TFP is defined as: $TFP_t = \frac{Y_t}{K_t^{\alpha} L_t^{1-\alpha}}$, where Y_t is GDP, K_t is the capital stock, L_t is hours worked, and $0 < \alpha < 1$ is the capital income share. For both countries we set a value of $\alpha = 0.3$. Following Conesa et al. (2007), we constructed the series for capital using the perpetual inventory method.

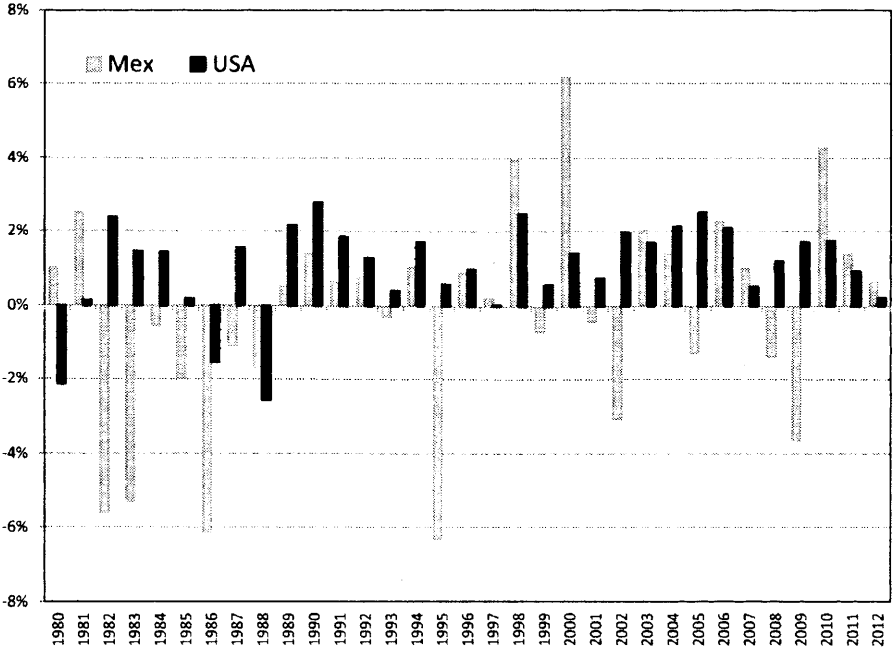


Figure I.6. TFP Growth (1980–2012): Mexico vs. U.S.
Source: Author's own calculations

particular, these authors show that a moderate negative TFP shock can trigger a bigger economic problem. The mechanism works as follows. If the negative shock results in a default by the government, then this causes a collapse of credit to the private sector. Consequently, producers, unable to use trade credit to purchase intermediate imported inputs, are forced to substitute towards domestic inputs. Economic efficiency is thereby reduced because imported and domestic inputs are not perfect substitutes, and in order to increase the supply of domestic inputs, labor is reallocated away from the production of the final good.

A related argument is presented in Kim (2014) who considers a model where TFP is endogenous and responds to shocks in the price of imported intermediate inputs. Similar to Mendoza and Yue (2012), he argues that imported and domestic inputs are imperfect substitutes. Furthermore, domestic inputs are less productive than imported ones, and therefore, overall efficiency can be reduced.

The study of business cycle fluctuations and economic crises in developing countries is a very important topic in the *macroeconomics of developing countries*. It is hoped that Chapter II of this book, by giving the reader an excellent introduction to this literature, will inspire future research in this area.

I.3.2 The informal sector

Another major divergence between developed and developing economies relates to differences in the institutional framework. Institutions shape the scope and extent of policies. Policies, in turn, distort the decisions taken by economic agents, which, at the end, prevent wealth creation (Parente and Prescott, 1994; Hayashi and Prescott, 2008). In developing countries the institutional framework is dysfunctional. Institutions and policies act as "barriers to riches" (Parente and Prescott, 2000). These barriers include stringent regulation, corruption, violence, and a lack of property rights, among others. A remarkable manifestation of this dysfunctional institutional framework is the importance of the informal sector in poor and middle-income countries. Informality is the result of the under-performance of tax collection and regulatory agencies in these countries. Figure I.7 shows the correlation between GDP per capita and the informality rate (a measure of the size of the informal sector) calculated by Schneider (2012). For the poorest countries in the sample, *the shadow economy* is as high as 65% of GDP, while for middle income countries it is around 30% of GDP. In Chapter IV of this book, Julio César Leal Ordóñez summarizes the informality literature with particular emphasis on the macroeconomic implications of the distortions arising from informality. One important lesson to emerge from this literature is the notion that the informal sector is not only a *symptom* of underdevelopment; it is also an important *cause* of it.

One highly debated area of research is the interaction between informality, financial frictions, and productivity. Some authors argue that credit constraints suffered by informal entrepreneurs lead to the misallocation of capital and low productivity. However, Midrigan and Xu (2013) and Moll (2014) have recently shown that the role of self-financing can help prevent this misallocation of capital. The reason is the following. If the idiosyncratic productivity shock enjoyed by an entrepreneur lasts for a sufficiently long period, then the entrepreneur will be able to achieve their optimal capital level by accumulating resources over time. Thus, capital misallocation will vanish in the long-run. This is an excellent example of why one has to be particularly cautious about the importance of financial frictions. If financial frictions are important for development in general, it has to be through a channel different from the misallocation of capital. One possibility (sketched by Jones, 2011) is that financial frictions create a *misallocation of ideas* by preventing optimal investment in technology innovations.

One aspect of informality that to date has received little attention relates to the variation of informality over the business cycle. Figure I.8 presents the cyclical components of output and the share of informal workers for the period 1993–2013.

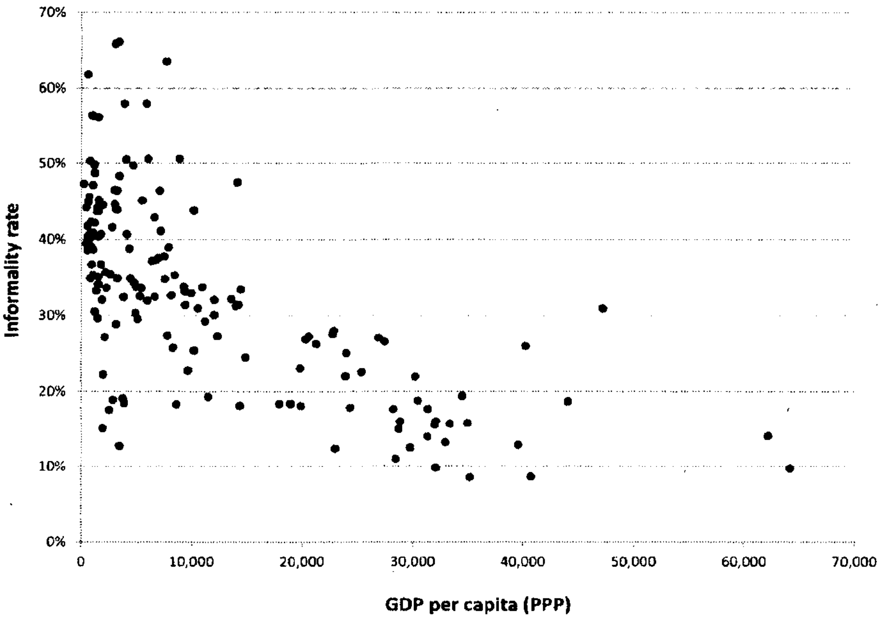


Figure I.7. Informality and GDP per capita
Source: Schneider (2012) and IMF-WEO

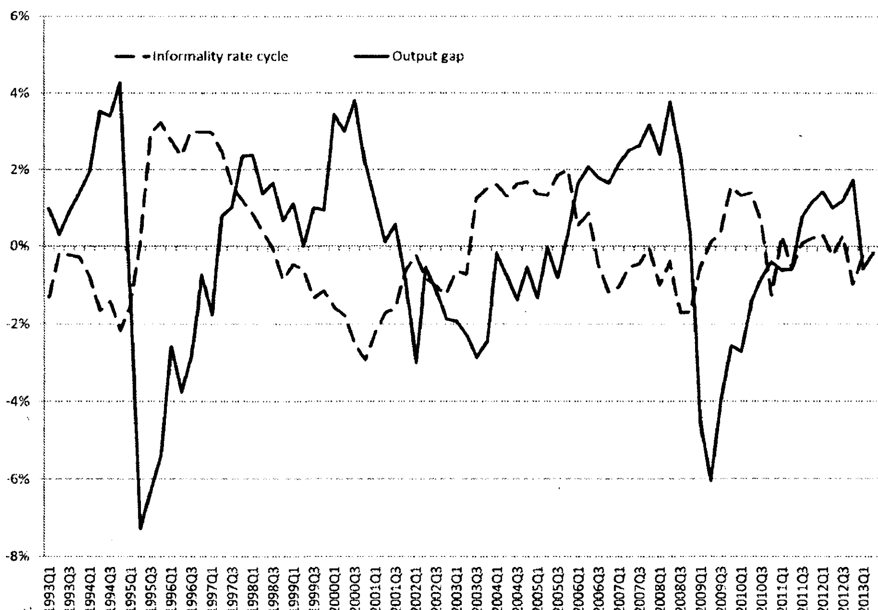


Figure I.8. Informality and the output gap
Source: Author's own calculations using INEGI data

Figure I.8 shows a strong negative correlation between the two series: informality goes up during output contractions, and decreases during expansions. Thus, an important research question relates to the causality of this strong negative correlation. Is it possible that the causality of this negative correlation can go in both directions? In other words, can an inflow of participants into the informal sector result in a significant downturn in aggregate output for the economy? Can the effects on output of a negative TFP shock be amplified through the expansion of the informal sector?

The study of the aggregate effects of informality (in both the short-run and long-run) is a rapidly expanding field. The major advancements made to date has been fueled by the use of firm heterogeneity models. Chapter IV, by summarizing this literature, aims to both inform and motivate readers on the importance of informality for the macroeconomy.

I.4 A BALANCED VIEW OF MACROECONOMICS

In this book we offer a balanced view of contemporary macroeconomics. Given the recent financial crisis, we understand the importance of studying this topic, and the ongoing policy issues and debates associated with it. We also believe that it is important to emphasize academic work oriented to understand the main problems faced by many developing countries, and that in our view, this research can also be informative to the problems facing the developed world. Of course, developed and developing countries also face many similar economic problems, particularly in issues relating to the design of monetary and fiscal policy. This is the final topic we now address.

I.4.1 Monetary policy for small open economies

A common feature of many developed and developing countries is that they are small in economic size. Table I.2 summarizes the degree of trade openness measured by the share of exports (or imports) in GDP for a number of developed and developing countries. As highlighted in Table I.2, small economies tend to have high ratios of international trade. Indeed, with the exception of the U.S., Japan, Australia, and Turkey, trade accounts for at least a quarter of total GDP (as measured by exports) over the period 2001–2012. However, despite the fact that many developed and developing countries are highly dependent on international trade, until recently the literature on monetary policy design had been conducted in the context of a closed economy. The general consensus of the closed-economy litera-

Table I.2. Mean exports and imports of goods and services: Selected countries (2001–2012)

2001–2012	*Mean exports (% of GDP)*	*Mean imports (% of GDP)*
	Developed countries	
Australia	19.8	21.1
Canada	35.4	33.7
France	26.5	27.4
Germany	43.3	38.2
Ireland	90.4	75.1
Japan	14.2	13.5
Netherlands	72.7	65.2
South Korea	44.5	42.3
Spain	27.4	30.7
Switzerland	49.0	40.5
United Kingdom	28.1	30.4
United States	11.0	15.3
	Developing countries	
Chile	38.2	32.1
China	31.1	26.7
Estonia	75.5	78.6
Hungary	76.3	75.3
Indonesia	30.8	26.2
Mexico	28.0	29.6
Poland	38.2	40.1
Russian Federation	32.4	22.2
South Africa	29.5	29.8
Turkey	23.7	26.8

Source: Author's own calculations using OECD data.

ture is that price stability could be achieved via the adoption of inflation-targeting policies by central banks (De Fiore and Liu, 2005). As discussed by Huang and Meng (2007), this trend began in industrial and middle-income countries in the late 1980s, and spread to developing and emerging economies in the late 1990s.

The best known example of an inflation-targeting policy is the celebrated Taylor (1993) rule:

$$i_t = \bar{r} + \pi_t + \mu_\pi \left(\pi_t - \pi^*\right) + \mu_y \left(y_t - \bar{y}\right), \tag{I.1}$$

where i_t denotes the nominal interest rate; $\bar{r}$ is the steady-state real interest rate; π_t is the inflation rate; π^* is the inflation target; $y_t - \bar{y}$ is the output gap; and $\mu_\pi > 0$, $\mu_y \geq 0$. If the inflation rate is above target, the policy rule requires that the central bank increase the nominal interest rate by more that the rate of inflation so that the real interest rate rises —the so-called *Taylor principle*. An important issue for small open-economies is whether policy rules like (I.1) and the Taylor principle are appropriate. In Chapter III of this book, Stephen McKnight summarizes the major findings from the open-economy literature. He outlines the conditions under which the Taylor principle may not be appropriate for open economies, and how the policy rule can be designed to ensure macroeconomic stability. It is hoped that this chapter will disseminate some of the important policy recommendations of this literature to a wider audience.

I.4.2 Taxes, transfers, and the labor market

Another common feature of both developed and developing countries is the need for governments' to raise revenue using taxation to provide a variety of public goods and services. In particular, one important strand of the macroeconomic literature is the study of taxation and its effects on aggregate outcomes. Spurred by the work of Prescott (2004), a recent branch of the literature seeks to determine the effect of taxes on labor supply. More generally, this literature looks for the answer to how sensitive is the aggregate labor supply to labor market shocks. While the traditional micro-labor literature reports that the labor supply elasticity is small, the labor supply elasticity assumed in both real business cycle and New Keynesian models is typically large. Thus, investigating the actual value of this elasticity is crucial for a range of macroeconomic topics. An important contribution in this literature is Rogerson and Wallenius (2009) who clarify the reasons behind the discrepancy between the micro and macro literatures.

Differences in income per-capita, labor market institutions, and in the configuration of taxes and transfers across OECD countries, provide rich information that is useful to study the aggregate labor supply elasticity. This is the main focus of Chapter VI of this book by Jorge Alonso-Ortiz. He shows that differences in labor market outcomes across OECD countries are greatly influenced by policy, and in particular, by differences in taxes and transfers. For example, Figure I.9 illustrates the relationship between annual hours worked per worker and

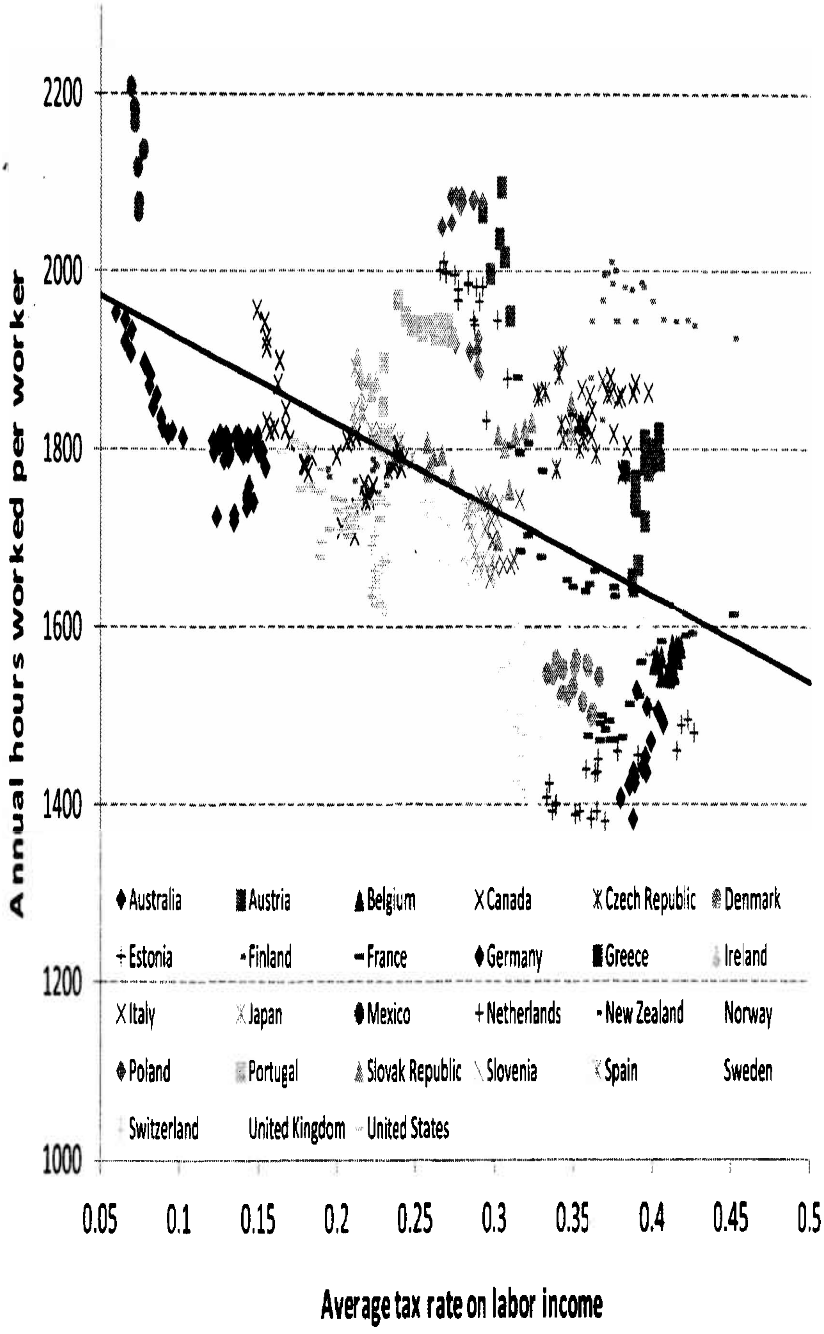


Figure I.9. Labor income taxes and hours per work

Source: McDaniel (2007) and Groningen Growth and Development Center average labor income taxes for 27 OECD countries for available time-periods.[10] Figure I.9 shows evidence of a clear negative correlation not only across countries, but also, in many cases, within countries too. The more general policy implications of these results are quite clear. Since the impact of differences in taxes and transfers on aggregate outcomes is sizable, countries can greatly improve income and welfare through appropriate policy changes.

It is hoped that Chapter VI of this book, by introducing the reader to this important literature, will help to both inspire future research and enlighten policymakers to implement better policies.

I.5 Organization of the Book

The rest of the book proceeds as follows.

Chapter II, by Arturo Antón Sarabia, provides an up-to-date summary on the study of business cycles. After outlining the key business cycle facts, with particular emphasis on the major differences between developed and developing economies, Antón presents a closed-economy neoclassical growth model with four possible distortions or wedges: efficiency, labor, investment, and government consumption. Employing the business cycle accounting method of Chari et al. (2007), the wedges are estimated using Mexican data, and the model is simulated to quantitatively assess the contribution of each wedge in explaining output fluctuations. The exercise for Mexico replicates the general findings of the literature, namely that the efficiency and labor wedges are the most promising distortions to account for output fluctuations. The author then outlines two well-known models that introduce frictions in the form of an efficiency wedge (Aguiar and Gopinath, 2007) and a labor wedge (Neumeyer and Perri, 2005), and explores how these models explain the observed differences between developed and developing countries at business cycle frequencies. The chapter concludes with a discussion on recent developments in the literature.

Chapter III, by Stephen McKnight, studies the design of monetary policy rules for small open economies. As discussed by the author, one important issue in the adoption of an interest rate rule by a central bank is that the policy rule should

[10] The initial and end time-period considered varies across countries. Australia (1959–2008), Austria (1995–2010), Belgium (1995–2010), Canada (1970–2010), Czech Republic (1995–2010), Denmark (1995–2010), Estonia (1995–2010), Finland (1975–2010), France (1950–2010), Germany (1995–2010), Greece (2005–2010), Ireland (2002–2010), Italy (1990–2010), Japan (1996–2008), Mexico (2003–2010), Netherlands (1990–2010), New Zealand (1986–2000), Norway (1978–2010), Poland (1989–2010), Portugal (1995–2010), Slovak Republic (1995–2010), Slovenia (1995–2010), Spain (2000–2010), Sweden (1995–2010), Switzerland (1995–2009), U.K. (1990–2010), U.S. (1970–2010).

avoid generating equilibrium indeterminacy, which can destabilize the economy through the emergence of welfare-reducing expectations-driven fluctuations. After outlining a small open economy model based on Gali and Monacelli (2005) and Faia and Monacelli (2008), McKnight first shows that economies open to international trade are more susceptible to indeterminacy than closed economies. This arises because interest rate changes generate adjustments in the terms of trade, which by altering inflation enables self-fulfilling expectations to be supported. Potential remedies to this indeterminacy problem and the appropriateness of managing the exchange rate are subsequently discussed. McKnight then extends the baseline model to include money, capital, and investment spending, and the chapter concludes by identifying the conditions under which it is desirable to use consumer price inflation and producer price inflation for the setting of interest rates.

Chapter IV, by Julio César Leal Ordóñez, considers the interaction between the informal sector and the aggregate economy. The author focuses on an important literature that analyzes the informal sector using dynamic general equilibrium models where informality plays a distortionary role. He first identifies the key ingredients in the informal sector literature: firm heterogeneity, and a net gain of informality that decreases with firm size. Then, in the context of these models, Leal discusses how informality can induce a number of distortions. For example, informal firms may choose to operate at a low scale in order to remain undetected by the government and continue to enjoy the benefits of tax avoidance. Informality is also shown to be related to resource misallocation across heterogeneous plants because it constitutes an implicit subsidy for low-productive firms. The author uses the model of Rauch (1991) that assumes discrete occupational choice and firm heterogeneity to present and discuss the distortions introduced into the economy with informality. He then considers several extensions to the baseline model that include: capital accumulation, own-account workers, heterogeneous skills, and financial constraints. The chapter concludes with a discussion of the main challenges facing the literature.

Chapter V, by Julio A. Carrillo and Céline Poilly, investigates the effects of different types of financial shocks on the real economy in the presence of near-zero interest rates. The authors develop a New Keynesian model enriched with the financial accelerator of Bernanke et al. (1999), and impose a zero lower bound constraint on the nominal interest rate. The effects of three financial shocks are considered: a net-worth shock that reduces the collateral value of borrowers, a risk shock that increases the uncertainty of the venture projects of borrowers, and a credit-spread shock that increases the cost of credit regardless of the financial health of borrowers. The authors find that only the credit-spread shock can push the nominal interest rate towards its zero floor causing a liquidity trap. It is then shown that loans are only procyclical under a risk shock. The other two shocks imply a counter-intuitive counter-cyclical response for loans. The chapter concludes by considering the effectiveness of a forward guidance monetary policy after a financial turmoil. The authors show that there is an optimal commitment period for the central bank to keep the nominal interest rate at the zero lower bound. However, beyond that optimal period, a forward guidance policy destabilizes the economy very quickly.

Chapter VI, by Jorge Alonso-Ortiz, studies the aggregate effects of changes in taxes and transfers. The author starts with a collection of twelve stylized facts on taxes, transfers, and aggregate hours using OECD data, stressing that there exists a negative correlation between hours of work per person and the labor tax wedge. Next, he presents a body of theory to help us understand this negative correlation. Starting with a static general equilibrium model, where the basic distortionary effect of labor taxes on aggregate hours can be studied, Alonso-Ortiz then proceeds to intertemporal models. He presents the findings of Prescott (2004), which utilizes a real business cycle model with capital, and then outlines the effect of taxes and transfers in models with uncertainty and incomplete markets. Key in this discussion is the way in which the main channels are affected by the presence of heterogeneous agents that are risk averse. Finally, since many transfers depend on age, the author presents a life-cycle model based on Erosa et al. (2012) where different components of social security can be studied. He concludes that the model is able to explain a substantial amount of observed cross-country differences in the labor supply patterns of people aged 50 and older.

The five papers in this volume have presented existing and on-going research in a number of important areas in macroeconomics. It is hoped that readers will find the book useful, not only as a learning device, but also to stimulate future research in these topics both here in Mexico and beyond.

REFERENCES

Acemoglu, D. (2009). The Crisis of 2008: Structural Lessons for and from Economics. Mimeo, MIT.

Aguiar, M., Gopinath, G. (2007). Emerging Market Business Cycles: The Cycle is the Trend. *Journal of Political Economy* 115, 69–102.

Arellano, C., Mendoza, E. (2002). Credit Frictions and Sudden Stops in Small Open Economies: An Equilibrium Business Cycle Framework for Emerging Markets Crises. NBER Working Paper No. 8880. National Bureau of Economic Research.

Bernanke, B., Gertler, M., Gilchrist, S. (1999). The Financial Accelerator in a Quantitative Business Cycle Framework. In Taylor, J.B., Woodford, M., eds., *Handbook of Macroeconomics*, 1341–1393. Amsterdam: Elsevier North-Holland.

Blanchard, O., Dell'Ariccia, G., Mauro, P. (2010). Rethinking Macroeconomic Policy. *Journal of Money, Credit and Banking* 42, 199–215.

Caballero, R. J. (2010). Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome. *Journal of Economic Perspectives* 24, 85–102.

Chari, V. V., Kehoe, P. J., McGrattan, E. R. (2007). Business Cycle Accounting. *Econometrica* 75, 781–836.

Carlstrom, C. T., Fuerst, T. S. (1997). Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. *American Economic Review* 87, 893–910.

Conesa, J. C., Kehoe, T., Ruhl, K. J. (2007). Modeling Great Depressions: The Depression in Finland in the 1990s. *Federal Reserve Bank of Minneapolis Quarterly Review* 31, 16–44.

De Fiore, F., Lui, Z. (2005). Does Trade Openness Matter for Aggregate Instability? *Journal of Economic Dynamics and Control* 29, 1165–1192.

Diamond, D. W., Dybvig, P. H. (1983). Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy* 91, 401–419.

Erosa, A., Fuster, L. Kamburov, G. (2012). Labor Supply and Government Programs: A Cross-Country Analysis. *Journal of Monetary Economics* 59, 84–107.

Faia, E., Monacelli, T. (2008). Optimal Monetary Policy in a Small Open Economy with Home Bias. *Journal of Money, Credit and Banking* 40, 721–750.

Gali, J., Monacelli, T. (2005). Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies* 72, 707–734.

García-Cicco, J., Pancrazi, R., Uribe, M. (2010). Real Business Cycles in Emerging Countries. *American Economic Review* 100, 2510–2531.

Gertler, M., Kiyotaki, N. (2010). Financial Intermediation and Credit Policy in Business Cycle Analysis. In Friedman, B. M., Woodford, M., eds., *Handbook of Monetary Economics*, 547–599. Amsterdam: Elsevier North-Holland.

Hayashi, F., Prescott, E. C. (2008). The Depressing Effect of Agricultural Institutions on the Prewar Japanese Economy. *Journal of Political Economy* 116, 573–632.

Huang, K. X. D., Meng, Q. (2007). Capital and Macroeconomic Instability in a Discrete-Time Model with Forward-Looking Interest Rate Rules. *Journal of Economic Dynamics and Control* 31, 2802–2826.

Jones, C. I. (2011). Misallocation, Economic Growth, and Input-Output Economics. NBER Working Papers No. 16742. National Bureau of Economic Research.

Kareken, J. H., Wallace, N. (1978). Deposit Insurance and Bank Regulation: A Partial-Equilibrium Exposition. *Journal of Business* 51, 413–438.

Kim, S. T. (2014). The Price of Imports and TFP: Application to the Korean Crisis of 1997–98. *Review of Economic Dynamics* 17, 39–51.

Kiyotaki, N., Moore, J. (1997). Credit Cycles. *Journal of Political Economy* 105, 211–248.

Krugman, P. (2009). How did Economists get it so Wrong? The New York Times, September 6th, 2009.

Kydland, F., Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica* 50, 1345–1370.

Kydland , F., Zarazaga, C. (2002). Argentina's Lost Decade. *Review of Economic Dynamics* 5, 152–165.

Reichlin, L., Baldwin, R., eds., (2013). *Is Inflation Targeting Dead? Central Banking After the Crisis.* London: Centre for Economic Policy Research.

McDaniel, C. (2007). Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD 1950-2003. Mimeo, Arizona State University.

McGrattan E. R., Prescott, E. C. (2012). The Great Recession and Delayed Economic Recovery: A Labor Productivity Puzzle? In Ohanian, L. E., Taylor, J. B., Wright, I.J., eds., *Government Policies and the Delayed Economic Recovery*, 115–154. Stanford, CA: Hoover Institution Press.

Mendoza, E. G., Yue, V. Z. (2012). A General Equilibrium Model of Sovereign Default and Business Cycles. *Quarterly Journal of Economics* 127, 889–946.

Midrigan, V., Xu, D. Y. (2013). Finance and Misallocation: Evidence From Plant-level Data. Forthcoming, *American Economic Review*.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? Forthcoming, *American Economic Review*.

Neumeyer, P. A., Perri, F. (2005). Business Cycles in Emerging Economies: The Role of Interest Rates. *Journal of Monetary Economics* 52, 345–380.

Parente, S. L., Prescott, E. C. (1994). Barriers to Technology Adoption and Development. *Journal of Political Economy* 102, 298–321.

Parente, S. L., Prescott, E. C. (2000). *Barriers to Riches*. Cambridge, MA: MIT Press.

Prescott, E. C. (2004). Why do Americans Work so Much More than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2–13.

Rauch, J. E. (1991). Modeling the Informal Sector Formally. *Journal of Development Economics* 35, 33–47.

Rogerson, R., Wallenius, J. (2009). Micro and Macro Elasticities in a Life-Cycle Model with Taxes. *Journal of Economic Theory* 144, 2277–2292.

Rogoff, K. (2010). Three Challenges Facing Modern Macroeconomics. *American Economic Association*. AEA Grand Challenges White Papers for Future Research.

Sargent, T. J. (2010). Modern Macroeconomics under Attack. Federal Reserve Bank of Minneapolis: *The Region*, September, 27–39.

Schneider, F. (2012). The Shadow Economy and Work in the Shadow: What do we (Not) Know? IZA Discussion Papers No. 6423. Institute for the Study of Labor.

Stiglitz, J. E. (2011). Rethinking Macroeconomics: What Failed, and How to Repair it. *Journal of the European Economic Association* 9, 591–645.

Taylor, J. B. (1993). Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.

Vergara, R. (2013). New Frameworks for Monetary Policy Analysis in an Era of Crises —Opening Remarks. *International Journal of Central Banking* 9, 3–11.

Woodford, M. (2009). Convergence in Macroeconomics: Elements of the New Synthesis. *American Economics Journal: Macroeconomics* 1, 267–279.

Woodford, M. (2012). What's Wrong with Economic Models? Research Note #009, Institute for New Economic Thinking.

Woodford, M. (2013). Inflation Targeting: Fix it, Don't Scrap it. In Reichlin, L., Baldwin, R., eds., *Is Inflation Targeting Dead? Central Banking After the Crisis,* 74–89. London: Centre for Economic Policy Research.

II. BUSINESS CYCLES IN DEVELOPED AND DEVELOPING COUNTRIES

ARTURO ANTÓN SARABIA
Centro de Investigación y Docencia Económicas

II.1 INTRODUCTION

Since the seminal work of Kydland and Prescott (1982), dynamic stochastic general equilibrium (DSGE) models have become the paradigm to understand and interpret business cycles. These models, underpinned by microeconomic foundations, require the researcher to be specific about the preferences, technologies, and constraints faced by every agent in the economy. As is well known, these class of models were developed as a response to the failure of the Keynesian paradigm in the 1970s to provide empirically correct answers, in particular to questions involving changes in monetary and fiscal policy. In a very influential paper, Lucas (1976) provided a series of examples in which expectations about future policy actions influence the current decisions of agents, and thus modify the empirical relationships in the data exploited by Keynesian model-builders at that time. Therefore, the way agents expectations are formulated must be explicit and consistent with individual maximization; they cannot be specified in an arbitrary manner. The so-called *Lucas critique* thus laid the foundations for the development of the modern macroeconomic models now widely used today. As forcefully stressed by Plosser (1989), the old view that the Keynesian paradigm was a success even if it lacked sound theoretical foundations could no longer be taken seriously.

The goal of this chapter is to present a series of business cycle facts in developed and developing economies, and to discuss how these facts are explained by representative macroeconomic models. As detailed below, while some of these facts are common to both developed and developing countries, others are not. Indeed, these differences constitute a serious challenge to macroeconomists if we require a single model to be able to explain such differences. In Section II.2, I begin by documenting a series of stylized facts for developed and developing countries at business cycle frequencies. This is an important first step, as the characterization of these facts are typically used to validate theoretical models.

Section II.3 presents the business cycle accounting method of Chari et al. (2007), which takes as a reference a neoclassical growth model with distortions or *wedges*. This method is designed in such a way that all wedges account for macroeconomic fluctuations in the model. This is an appealing methodology since a large class of frictions in DSGE models may be represented in terms of four wedges. For this reason, it provides a useful guide for researchers as to what class of distortions may allow DSGE models to account for output fluctuations in the data. For illustrative purposes, the method is applied to Mexican data. The exercise suggests that two wedges, the efficiency and labor wedges, are the most promising to explain macroeconomic fluctuations in Mexico.[1] Indeed, similar results have been found for a number of other developing and developed economies.[2] Therefore, these studies suggest that any model that aims at understanding macroeconomic fluctuations in *all* countries must incorporate frictions of this nature.

In Section II.4, I then discuss a number of DSGE models that are commonly used to understand business cycle differences in developed and developing economies. The discussion centers around the two well-known models of Aguiar and Gopinath (2007), and Neumeyer and Perri (2005). From the lens of the business cycle accounting method, Aguiar and Gopinath (2007) introduce efficiency wedge frictions by decomposing productivity into a transitory and a permanent component. In their model, the volatility of permanent productivity shocks is crucial to explain the differences between developed and developing countries. On the other hand, frictions in the Neumeyer and Perri (2005) model take the form of a labor wedge. In their framework, firms need to pay labor services before production takes place, thus creating a need for working capital. Given that firms must borrow to cover labor payments, labor demand depends inversely on the real interest rate. If labor supply is independent of consumption, increases in the real interest rate decrease equilibrium employment and thus output. This mechanism explains the negative correlation between output and interest rates found in the data for developing economies.[3] Finally, Section II.5 briefly concludes.

[1] The efficiency wedge represents distortions between factor inputs and output, whereas the labor wedge represents a gap in the marginal rate of substitution between leisure and consumption.

[2] See, e.g., Chari et al. (2007) for the U.S.; Antón (2008) for Mexico; Kersting (2008) for the U.K.; Lama (2011) for several Latin American countries.

[3] In addition, Neumeyer and Perri (2005) show that labor wedge frictions can also explain a number of other business cycle properties for developing countries.

II.2 COMPARING BUSINESS CYCLES IN DEVELOPED AND DEVELOPING ECONOMIES

This section provides a discussion on both the similarities and differences between developed and developing economies at business cycle frequencies.[4] First, I outline alternative definitions of the business cycle. Next, business cycles in developed and developing countries are compared in terms of their duration and amplitude. Finally, I identify the key stylized facts, in terms of moments in the data, typically used in the literature to validate theoretical models.

II.2.1 Defining business cycles

There is no consensus among economists on how to define a business cycle. Roughly speaking, there are two alternative definitions. The *classical* view, provided by Burns and Mitchell (1946), defines a business cycle as a series of expansions that occur about the same time in many economic activities, followed by similarly general contractions and revivals, which merge into the expansion phase of the next cycle. In contrast, the *growth* definition of business cycles refers to recurrent fluctuations of GDP around a trend, and the co-movements among other aggregate time series (Lucas, 1977).

Each definition requires the employment of different methods in order to characterize a business cycle. Typically, the classical approach for the identification of a business cycle is based on the Bry and Boschan (1971) algorithm, or its extended version (Harding and Pagan, 2002). This method allows the researcher to identify the peaks and troughs along a cycle (i.e. the turning points), as well as its duration and amplitude. On the other hand, the growth definition generally requires the researcher to decompose a (seasonally adjusted) time series into a cycle and a trend component.[5] Once the cyclical component of the series is obtained, the business cycle may be described in terms of relevant moments like standard deviations and correlations. Clearly, each method emphasizes different aspects of the business cycle: duration and amplitude under the classical approach; and standard deviations and correlations under the growth approach. For this reason, the information provided by each method should be seen as complementary to each other. However, it is the growth approach which is widely used in the literature to validate theoretical models.

[4] The following discussion concentrates on real variables, rather than nominal variables, given that this literature typically focuses on the real side of the economy.

[5] See Canova (2007) for a discussion on the alternative methods available to detrend time series.

II.2.2 Duration and amplitude

Business cycles in developed and developing countries differ substantially in terms of their duration. For the former, Artis et al. (1997) employ a version of the Bry-Boschan algorithm to identify business cycles for the G-7 and five additional developed European countries using monthly data on industrial production in each country for the period 1961(1)–1993(12). On average, they find that the duration of contractions and expansions are 15 and 51 months, respectively, implying an average business cycle of 66 months. On the other hand, Rand and Tarp (2002) apply the Bry-Boschan method for a sample of 15 developing countries using indexes of industrial production as a proxy for aggregate output. While the data period varies depending on the country, it roughly covers the period 1960–1999. They document that the average duration of contractions and expansions for the whole sample is 15.6 and 14.4 months, respectively, yielding an average duration of the cycle of 30 months. The combined evidence of these two studies suggest that while the duration of contractions is roughly similar in both developed and developing countries, expansions in the latter are relatively short-lived. Consequently, a complete cycle has a shorter duration in developing countries, implying that business cycles move relatively quickly from trough to peak.

However, a major weakness of these studies is that the results are generated using industrial production data. To the extent that industrial production at business cycle frequencies follows a different pattern than a more aggregate measure of economic activity (say, GDP), the inference made may be misleading. This concern was addressed by Calderón and Fuentes (2010) who employ the Harding and Pagan (2002) algorithm to identify a series of business cycle facts using GDP quarterly data. Specifically, they estimate the duration and amplitude of the business cycle for a sample of 23 emerging markets (12 Latin American countries, 8 East Asian and Pacific, and 3 other emerging economies), and 12 industrial economies over the period 1980–2006. Their results are summarized in Table II.1.

Consistent with the findings using industrial production data, Calderón and Fuentes (2010) report that the average duration of contractions ("Contrac.") is roughly similar across country groups, with an interval between 3.5 and 4.8 quarters. However, contrastingly, the duration of expansions ("Expan.") differ substantially across groups. For example, expansions last about 16 quarters on average for Latin American countries, which fall short from the 21 and 24 quarters on average for East Asian and developed countries, respectively. Nonetheless, the

Table II.1. Business cycles: duration and amplitude (quarterly data, 1980Q1–2006Q2)

	Mean duration (quarters)		*Mean amplitude (%)*		*Number of contractions*
	Contrac.	*Expan.*	*Contrac.*	*Expan.*	
Latin America	3.5	16.0	−6.2	21.3	4.8
Asia	4.2	21.3	−7.4	41.6	2.9
Other Emerging	4.8	17.1	−4.8	28.9	3.3
All Emerging	4.0	17.3	−6.6	27.9	4.1
OECD Countries	3.6	23.8	−2.2	20.2	3.3

Source: Calderón and Fuentes (2010).

general result that expansions are of shorter duration in developing countries relative to developed countries is maintained (17.3 versus 23.8 quarters).[6]

Calderón and Fuentes (2010) also find large differences in the amplitude of the cycle between industrial and emerging economies. For the case of OECD countries, the mean amplitude from peak to trough is only 2.2 percent, whereas in Latin American and East Asian countries such amplitude is 6.2 and 7.4 percent, respectively. This is an important observation: contractions last roughly the same in developed and developing countries, but for the latter the magnitude of the output fall is significantly larger. The amplitude of the expansion is also larger in developing countries, especially in Asian countries. These two pieces of evidence suggests that output fluctuations are smaller in industrial economies. The last column in Table II.1 also shows that Latin American countries experienced roughly five contractions during the sample period. In contrast, all other groups of countries had only three contractions during the same period. This implies that Latin America is a relatively volatile region, with sharp and frequent output falls.

II.2.3 Data moments

Since the work of Backus and Kehoe (1992), a series of business cycle stylized facts for developed economies have been identified. In their study, Backus and Kehoe (1992) use macroeconomic data for ten countries with a time-span of over a century. To remove low-frequency movements from the data (i.e., the trend

[6] For the particular case of Mexico, Antón (2011) also reports that the duration of contractions is similar to those registered in developed countries, but the duration of expansions is significantly shorter.

component of the series), they employed the commonly used Hodrick-Prescott (HP, 1997) method. From this perspective, output fluctuations are understood as the standard deviation of the HP filtered logarithm of real output. Similarly, co-movements between GDP and macroeconomic variables are measured in terms of correlations. Despite the differences in institutions, and fiscal and monetary policies, Backus and Kehoe (1992) report a series of empirical regularities for the sample countries. For example, investment is about 2–4 times as volatile as output; consumption is as volatile as output; investment and consumption are strongly procyclical; and the trade balance is slightly countercyclical, meaning that it exhibits larger deficits during booms than during recessions.

Some of these regularities found by Backus and Kehoe (1992) have also been reported for developing countries. Rand and Tarp (2002) use data for 15 developing economies (5 from Sub-Saharan Africa, 5 from Latin America, and 5 from Asia and North Africa) for the period 1970–1997. To detrend the data, they use the HP filter, as well as the band-pass (BP) filter originally proposed by Baxter and King (1999). They find that investment is more volatile than output; and both total consumption and investment are procyclical (see also Mendoza, 1995). These properties are in line with those observed in OECD countries. In terms of output persistence, the evidence suggests that it is high, with no significant differences between developed and developing countries in general (see, e.g., Backus et al., 1995; Mendoza, 1995). However, despite some common regularities, there are also major differences between developed and developing economies.

Stylized difference 1. Mendoza (1995) reports a higher output volatility in developing countries on average. In particular, the mean standard deviation of output is 1.5 times larger in developing countries. From all the sample groups analyzed, the Latin American region is the group with the highest average volatility (1.9 times larger than G-7 countries).[7] Notice that these results complement the findings from the cyclical approach reported above. The fact that output volatility is relatively higher in developing countries has been reported in a number of other studies.[8] Table II.2 summarizes the moments at business cycle frequencies for developed and emerging economies as reported by Neumeyer and Perri (2005). Panel A shows that the standard deviation of output is about 2.8 percent in emerging economies, but only about 1.4 percent in developed countries.

[7] There is evidence that a higher volatility in Latin American countries has also been present before World War II. Aiolfi et al. (2011) use econometric techniques to construct a business cycle index for Argentina, Brazil, Chile, and Mexico that spans for 135 years. They find evidence that these four countries experienced significantly greater output volatility than advanced economies during the period 1870–1929.

[8] See, e.g., Rand and Tarp (2002), Neumeyer and Perri (2005), Talvi and Végh (2005), Aguiar and Gopinath (2007).

Table II.2. Business cycles in emerging and developed economies

	(A) Standard deviations							
	GDP	RIR	NX	PC	TC	INV	EMP	HRS
Emerging	2.79	2.32	2.40	1.30	1.71	3.29	0.73	0.89
Developed	1.37	1.66	0.92	0.92	1.08	3.44	1.11	1.61
	(B) Correlations with GDP							
		RIR	NX	PC	TC	INV	EMP	HRS
Emerging	–	−0.55	−0.61	0.80	0.79	0.88	0.54	0.65
Developed	–	0.20	−0.23	0.67	0.68	0.73	0.81	0.84
	(C) Correlations with interest rate							
			NX	PC	TC	INV	EMP	HRS
Emerging	–	–	0.51	−0.55	−0.56	−0.48	−0.53	−0.52
Developed	–	–	−0.22	0.24	0.25	0.21	0.30	0.11

Notes: The percentage standard deviations of all variables are divided by the percentage standard deviation of output, except for the variables GDP, NX and RIR, which are expressed as the percentage standard deviation only. Net exports (NX) are exports minus imports over GDP. Real interest rates (RIR) are in percentage points. Total consumption (TC) includes private consumption (PC) and government consumption, changes in inventories and a statistical discrepancy. Investment (INV) is gross fixed capital formation. Employment (EMP) is number of workers, and total hours (HRS) is number of workers times weekly hours of work per worker.

Source: Neumeyer and Perri (2005).

Stylized difference 2. Mendoza (1995) reports that total private consumption is relatively more volatile than GDP in developing countries, and less volatile than GDP in G-7 countries. This result has also been found by Rand and Tarp (2002), Neumeyer and Perri (2005), and Aguiar and Gopinath (2007), and it is illustrated in Panel A of Table II.2. The standard deviation of private consumption relative to that of output is 1.3 times larger in emerging economies, but it is roughly 10 percent smaller in developed countries. Furthermore, public consumption is also relatively more volatile than GDP in developing countries, than developed economies.

Stylized difference 3. Neumeyer and Perri (2005) report that real interest rates in developing economies are countercyclical and lead the business cycle.[9] In sharp contrast, real interest rates in developed economies are acyclical and lag the cycle. To illustrate this point, contemporary correlations for output and interest rates with relevant macroeconomic variables in emerging and developing economies are shown in Panels B and C of Table II.2. This finding implies that for developing economies, periods of low interest rates are generally associated with economic expansions, and periods of high interest rates are often characterized by low levels of economic activity (see also Uribe and Yue, 2006).

Stylized difference 4. There is evidence that the trade balance is relatively more volatile in developing countries. It is also mildly procyclical in G-7 countries, weakly countercyclical in small developed economies, and strongly countercyclical in developing economies (see, e.g., Mendoza, 1995; Neumeyer and Perri, 2005). These results can be seen by inspection of Panels A and B of Table II.2. On the other hand, Mendoza (1995) reports that the terms of trade are more volatile in developing economies than G-7 countries, both in absolute terms and relative to GDP. The terms of trade are also strongly procyclical in G-7 countries, but only mildly procyclical in developing countries on average. He also finds that real effective exchange rates are substantially more volatile in developing countries, but acyclical in both developed and developing economies.

Stylized difference 5. Talvi and Végh (2005), using a sample of 56 countries, of which 20 are industrial and 36 are developing economies for the period 1970–1994, find that government consumption is acyclical in G-7 countries, mildly procyclical in other industrial countries, and strongly procyclical in developing economies. In fact, none of the developing countries in the sample showed a negative correlation between government consumption and output. These differences in co-movements between developed and developing countries are also found in Rand and Tarp (2002).

II.3 A BENCHMARK METHOD FOR THE STUDY OF BUSINESS CYCLES

The stylized facts reported in the previous section constitute an essential ingredient for formulating and validating theoretical models. Since the seminal work of Kydland and Prescott (1982), DGSE models have become the dominant paradigm

[9] Here, real interest rates capture the cost of borrowing that economies face in international financial markets.

for the study of business cycles. One of the major advantages of these models is that they are built from microeconomic principles. This requires the researcher to specify the fundamentals of the economy, such as preferences and technology, as well as the constraints faced by all economic agents. By construction, this structure provides a clear mechanism of how exogenous shocks are transmitted through the economy.

Constructing and solving DSGE models to interpret business cycles is a major task. Typically, these models are simulated by assuming some exogenous process for the shocks. These shocks affect the endogenous variables of the model so that the relevant moments for the variables of interest may be computed. The moments from the artificial economy may then be compared to the corresponding moments in the data in order to assess how successful the model is in replicating observed business cycles. Of course, this task is even more daunting if a single model is used to try and understand the differences in stylized business cycle facts between developed and developing countries. In pursuing this task, researchers typically face difficult choices about where to introduce frictions into their DSGE models in order to generate business cycle fluctuations similar to those in the data. Recently, Chari et al. (2007) have proposed a relatively simple method to guide such choices: the so-called *business cycle accounting* approach. This method is appealing since many macroeconomic models of fluctuations are equivalent to the benchmark growth model with distortions as proposed by the authors.

The prototype model of Chari et al. (2007) described below is a neoclassical, closed-economy growth model, with four stochastic variables or wedges: efficiency, labor, investment, and government consumption. These time-varying wedges distort the equilibrium decisions of agents operating in otherwise competitive markets. The wedges are first estimated from both the data and the equilibrium conditions of the benchmark prototype economy, and then fed back into the model to quantitatively account for the contribution of the wedges to business cycle fluctuations, either separately or in combinations. By construction, the four wedges fully account for the observed movements in macroeconomic variables. Hence the name *business cycle accounting*. Below, I outline the benchmark closed-economy model of Chari et al. (2007).[10] The estimation method is then discussed briefly and applied to Mexican data for illustrative purposes.

[10] For an exposition of the method in a small open economy setting, see Lama (2011) for details.

II.3.1 The benchmark prototype economy

The prototype economy consists of a standard neoclassical growth model with adjustment costs for investment and four exogenous stochastic variables labeled as follows: the efficiency wedge A_t, the labor wedge $1-\tau_{n,t}$, the investment wedge $1/(1+\tilde{\tau}_{x,t})$, and the government consumption wedge g_t. Households maximize lifetime expected utility over per capita consumption c_t and per capita labor l_t streams:

$$\mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_t, l_t) N_t \tag{II.1}$$

subject to the budget constraint

$$c_t + (1+\tau_{x,t})x_t = (1-\tau_{n,t})w_t l_t + r_t k_t + T_t, \tag{II.2}$$

and the law of motion for capital

$$(1+\gamma_n)k_{t+1} = (1-\delta)k_t + x_t - \varphi\left(\frac{x_t}{k_t}\right)k_t, \tag{II.3}$$

given the initial condition $k_0 > 0$. In the expressions above, N_t is the period t population growing at the rate $1+\gamma_n$, x_t is investment, w_t is the wage rate, r_t is the rental rate of capital, k_t is the per capita capital stock, T_t is per capita lump-sum transfers, $\delta \in (0,1)$ is the depreciation rate of capital, $\beta \in (0,1)$ is the discount factor, and $\tau_{x,t}$ $(\tau_{n,t})$ is the tax rate on investment (labor). The function $\varphi\left(x_t/k_t\right)$ represents the costs of adjusting investment with properties $\varphi' > 0$ and $\varphi'' \geq 0$. Adjustment costs are incorporated into the model since Christiano and Davis (2006) find that the accounting exercise may be sensitive to the value of Tobin's q elasticity.[11]

Technology is represented by a neoclassical production function of the form $F\left(k_t, (1+\gamma)^t l_t\right)$, where the term $(1+\gamma)^t$ is the exogenous growth rate of labor-augmenting technical progress. Thus, per capita output y_t is determined by:

$$y_t = A_t F\left(k_t, (1+\gamma)^t l_t\right). \tag{II.4}$$

As is usual in a perfectly competitive environment, the price of each factor of production is equal to their corresponding marginal productivity, i.e., $w_t = F_{n,t}$ and $r_t = F_{k,t}$. Finally, government in the model is represented in terms of per

[11] Tobin's q elasticity ζ is defined as the elasticity of investment-to-capital ratio with respect to the price of capital. Namely, $\zeta \equiv \frac{d \log(x_t/k_t)}{d \log P_{k,t}}$, where $P_{k,t} = 1/\left(1-\varphi'(x_{t+1}/k_{t+1})\right)$ is the market price of capital in the benchmark prototype economy.

capita expenditures g_t, where g_t fluctuates around a trend given by $(1+\gamma)^t$. The resource constraint in this economy is thus given by:

$$c_t + x_t + g_t = y_t. \tag{II.5}$$

Standard first-order conditions of the household problem yield:

$$-\frac{U_{l,t}}{U_{c,t}} = \left(1 - \tau_{n,t}\right) A_t (1+\gamma)^t F_{n,t}, \tag{II.6}$$

$$(1 + \tilde{\tau}_{x,t}) U_{c,t} = \beta \mathrm{E}_t U_{c,t+1} \left[A_{t+1} F_{k,t+1} + \left(1 + \tilde{\tau}_{x,t+1}\right) \Phi_{t+1}\right], \tag{II.7}$$

where $U_{j,t}$ denotes the derivative of U_t with respect to $j = c, l$; $1 + \tilde{\tau}_{x,t} \equiv \frac{1+\tau_{x,t}}{1-\varphi'(x_t/k_t)}$; and

$$\Phi_{t+1} \equiv \left[1 - \delta - \varphi\left(\frac{x_{t+1}}{k_{t+1}}\right) + \varphi'\left(\frac{x_{t+1}}{k_{t+1}}\right)\left(\frac{x_{t+1}}{k_{t+1}}\right)\right].$$

The left hand side of equation (II.6) is the marginal rate of substitution between leisure and consumption, which is in turn equal to the after-tax marginal product of labor. Expression (II.7) is the familiar Euler equation, where intertemporal consumption is a function of the adjusted investment tax rate $\tilde{\tau}_{x,t}$.[12]

In this model, the efficiency wedge A_t in (II.4) resembles the productivity parameter. In a similar fashion, the terms $1 - \tau_{n,t}$ and $1/\left(1 + \tilde{\tau}_{x,t}\right)$ introduce a wedge in expressions (II.6) and (II.7) with respect to an otherwise standard neoclassical model with no distortions. These wedges resemble (but are not necessarily equal to) the tax rates on labor income and investment. Finally, the government consumption wedge g_t is defined by (II.5). This system of four equations solves for the four wedges in the model, given the law of motion for capital (II.3).

As discussed by Chari et al. (2007), the appeal of this relatively simple framework is that a large class of macroeconomic models may be mapped into the benchmark prototype economy described above. For example, a model with constant technology and input-financing frictions is equivalent to a growth model with efficiency wedges. Alternatively, an economy with sticky wages and monetary shocks is equivalent to a prototype model with labor wedges, and so on.

[12] This intertemporal wedge may alternatively be defined in terms of a tax on capital income $\tau_{k,t}$. Chari et al. (2006) find that the accounting procedure is not sensitive to this alternative specification.

II.3.2 Estimation method

The accounting procedure originally developed by Chari et al. (2007) is implemented in two steps. First, wedges of the benchmark prototype economy are measured by using both the data and a detrended (i.e., stationary) version of the model. Then the prototype model is simulated using the wedges already obtained to assess their contribution (either separately or in combinations) to fluctuations in variables of interest, such as output, labor, or investment.

Measurement of wedges. The measurement of wedges is performed in three steps.

1. First, functional forms for preferences and technology are set, and parameter values of the benchmark prototype model are calibrated in the spirit of the business cycle literature.[13] In our example, preferences are of the logarithmic form $U(c, l) = \log c + \psi \log(1 - l)$, the production function is Cobb-Douglas $F(k, l) = k^{\alpha} l^{1-\alpha}$, and the adjustment costs are specified in terms of a quadratic function $\varphi(x/k) = (a/2)(x/k - b)^2$.[14] In these functions, the parameters ψ, α, a, and b are calibrated so that the steady state properties of the model replicate relevant features of the data.

2. Second, the stochastic process for the wedges is estimated using the assumed functional forms, and the calibration of parameter values. For such estimation, a vector autoregressive VAR(1) process of the form is assumed:

$$s_{t+1} = P_0 + P s_t + \varepsilon_{t+1}. \tag{II.8}$$

   Here, the vector s_t is defined in terms of the four wedges, namely

$$s_t = (A_t, \tau_{n,t}, \tilde{\tau}_{x,t}, g_t),$$

   and the shock ε_t is i.i.d. and distributed normally with mean zero and covariance matrix V. The parameters included in matrices P_0, P, and V of the VAR(1) process for the wedges are then estimated using maximum likelihood methods, the log-linear decision rules of the benchmark prototype economy, and data on output, labor, investment, and government consumption including net exports (see Chari et al. 2006, for further details).

[13] See, e.g., Kydland and Prescott (1982), Prescott (1986), King et al. (1988).

[14] Chari et al. (2007) show that the business cycle accounting method is qualitatively robust to alternative specifications of production functions and preferences.

3. Once the stochastic process in (II.8) is estimated, the final step is to measure the four wedges. These wedges may be recovered from the data and the equilibrium conditions of the benchmark economy. For example, the government consumption wedge may be measured directly from the data as the sum of government expenditures and net exports so that the data is consistent with the theory. To measure the remaining three wedges let y_t^d, l_t^d, x_t^d, and k_0^d denote data on production, labor, investment, and the initial capital stock, respectively, and let $y(s_t, k_t)$, $l(s_t, k_t)$, and $x(s_t, k_t)$ represent the decision rules of the model. Then the realized wedge series s_t^d solves:

$$y_t^d = y(s_t^d, k_t), l_t^d = l(s_t^d, k_t), x_t^d = x(s_t^d, k_t), \quad \text{(II.9)}$$

with $k_{t+1} = (1 - \delta)k_t + x_t^d - \varphi\left(x_t^d/k_t\right)k_t$, $k_0 = k_0^d$, and $g_t = g_t^d$. Thus, the three equations (II.4), (II.6) and (II.7) are used to solve for the remaining three unknown elements of the vector s_t.

Contribution of wedges. After estimating the wedges, the benchmark prototype model may be simulated in order to assess, separately and in combinations, the contribution of the wedges to fluctuations in the variables of interest, starting at some initial date. This contribution is measured by comparing the realizations of variables such as output, labor, and investment that arise from simulating the model with those in the data. For example, consider defining the vector of wedges $s_{1,t} = (A_t, \bar{\tau}_n, \bar{\tau}_x, \bar{g})$ so that in period t the efficiency wedge takes its period t value while the other wedges are kept at some constant values. The corresponding decision rules for output, labor, and investment may be denoted by $y^e(s_{1,t}, k_t)$, $l^e(s_{1,t}, k_t)$, and $x^e(s_{1,t}, k_t)$, respectively. These decision rules along with an initial condition k_0^d, the realized wedge series s_t^d, and the law of motion for capital may be used to compute sequences for output, labor, and investment, denoted by y_t^e, l_t^e, and x_t^e, respectively. These sequences are called the *efficiency wedge components* of output, labor, and investment. These output, labor, and investment components can then directly be compared to the actual data to assess how well they can match the data.

Of course, this accounting exercise may be performed in alternative ways. For example, the labor wedge components may be computed in a similar fashion by defining a vector $s_{2,t} = (\bar{A}, \tau_{n,t}, \bar{\tau}_x, \bar{g})$, and so on. It is also possible to construct components for combined wedges. For example, the efficiency plus labor wedge components may be obtained after defining a vector $s_{5,t} = (A_t, \tau_{n,t}, \bar{\tau}_x, \bar{g})$. If the four wedges are then fed into the decision rules in (II.9), and used in combination with both the law of motion for capital and the equation $g_t(s_t^d) = g_t$, all the movements in output, labor, and investment from the simulation are exactly those observed in the data by construction.

II.3.3 Illustrative results

The accounting procedure of Chari et al. (2007) is now illustrated using quarterly data for Mexico. The period of study is 1987Q1–2006Q3. This period is interesting, as it includes both the so-called *Tequila crisis* of 1994–1995 and the 2001 recession. The recession of 1994–1995 was particularly sharp, as detrended output per capita fell slightly more than 12 percent, two quarters right after the beginning of the crisis. Statistical properties of the output components are presented in Table II.3, assuming an elasticity value for Tobin's q of 3.[15] The first entry in Part A shows that the efficiency wedge component of output fluctuates relatively less with respect to actual output, and it is positively correlated with contemporaneous output for several leads and lags. A similar qualitative result is found for output due to the labor wedge component, although it is able to explain a higher fraction of output volatility in the data. In contrast, both investment and the government consumption wedge component are negatively correlated with output. Part B of Table II.3 summarizes the cross correlations between the output components from two combined wedges. The results illustrate that output components from the efficiency and labor wedges are positively correlated for all leads and lags. In contrast, all the other possible combinations of output components are either essentially non-correlated or negatively correlated for all leads and lags.

The conclusions from Table II.3 are as follows. The efficiency wedge is the most promising friction in helping to explain output fluctuations in Mexico, since it exhibits the highest contemporaneous correlation with output combined with a substantial output variability. Similarly, the labor wedge is also promising although it displays a slightly lower correlation with output. The investment wedge is less important than the other two, since it only contributes in explaining less than 20 percent of observed output variability, and it is (slightly) negatively correlated with actual output. Finally, output due to the government consumption wedge is simply inconsistent with actual data, given its negative correlation with output.

Another way to evaluate these results is by simulating how much the efficiency and the labor wedge components can separately account for output falls during recessions. Figures II.1 and II.2 present data on detrended output, along with the

[15] The interested reader is referred to Antón (2008) for details on the estimation and calibration of parameters.

Table II.3. Properties of the output components with Tobin's q elasticity $q = 3$

(A) Summary statistics

Output components	*Standard deviations*	*Cross correlation of wedge with output at lag k =* −2	−1	0	1	2
Efficiency	0.76	0.48	0.74	0.91	0.74	0.51
Labor	1.00	0.53	0.72	0.78	0.71	0.59
Investment	0.22	−0.01	0.04	−0.04	−0.05	−0.11
Government	0.83	−0.33	−0.55	−0.62	−0.58	−0.52

(B) Cross correlations

Output components (X, Y)	*Cross correlation of X with Y at lag k =* −2	−1	0	1	2
Efficiency, Labor	0.45	0.61	0.61	0.59	0.44
Efficiency, Investment	−0.12	−0.10	−0.17	−0.02	−0.06
Efficiency, Government	−0.40	−0.51	−0.56	−0.46	−0.28
Labor, Investment	−0.02	0.06	0.06	0.01	0.08
Labor, Government	−0.55	−0.67	−0.91	−0.62	−0.40
Investment, Government	−0.11	−0.07	−0.19	−0.15	−0.04

Notes: The sample period is 1987Q1–2006Q3. The standard deviations are all relative to output.
Source: Antón (2008).

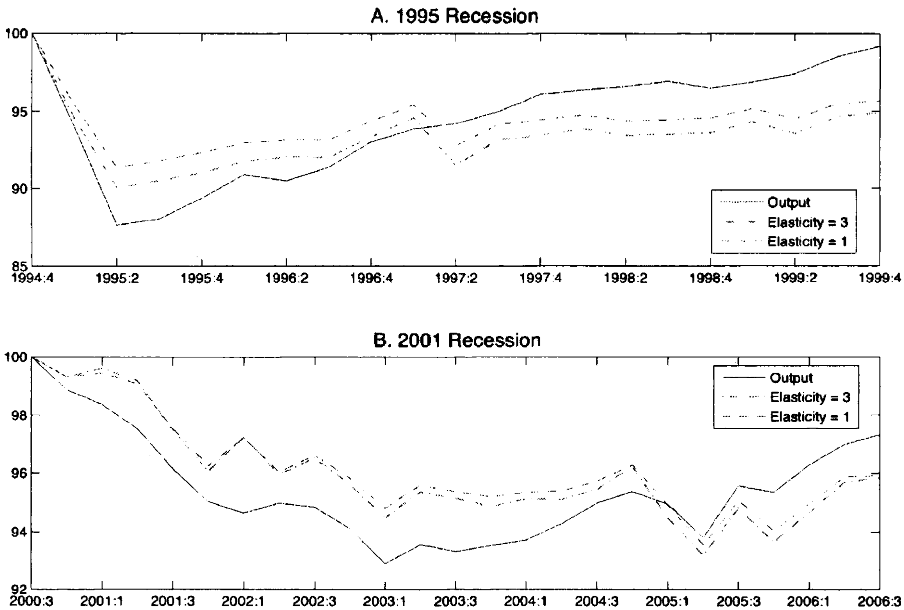


Figure II.1. Output & predictions of model with the efficiency wedge only

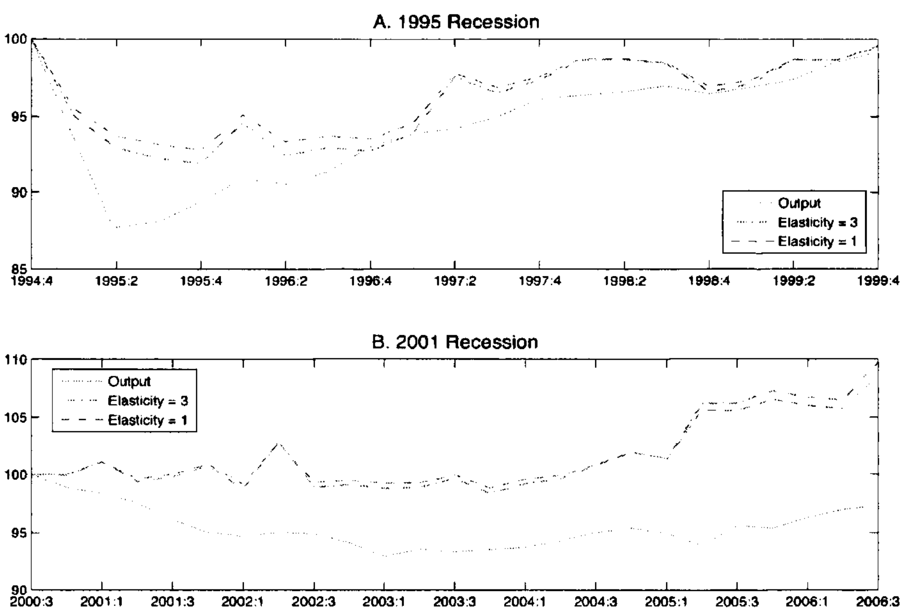


Figure II.2. Output & predictions of model with the labor wedge only predictions of the model due to the efficiency and labor wedge components, respectively, under alternative values of Tobin's q elasticity. The exercise is applied to the 1995 and 2001 recessions, where for illustrative purposes output is normalized to 100 at the beginning of each recession. Lets first consider Figure II.1. Consistent with the previous discussion, simulated output due to the efficiency wedge component follows the actual data relatively close regardless of the value for Tobin's q elasticity. In fact, this wedge roughly explains between 70 and 80 percent of the output fall at the trough of either the 1995 or the 2001 recession, depending on the parametrization of the adjustment cost function. Figure II.2 presents the results when only the labor wedge component is in effect. Depending on the value of Tobin's q elasticity, the labor wedge is able to explain between 51 and 58 percent of the output fall at the trough of the 1995 recession. It also captures the output recovery relatively well. However, the labor wedge component only accounts for between 11 and 17 percent of the output fall at the trough of the 2001 recession, and it is less correlated with actual data.

To summarize, this illustrative exercise suggests that the efficiency and labor wedges are the two most promising distortions accounting for output fluctuations in Mexico. In contrast, both the investment and government consumption wedges play a minor role in understanding fluctuations.[16] This general result has been confirmed in a series of studies.[17] Remarkably, these results are robust to alternative specifications of the benchmark economy (e.g., the inclusion of variable capital utilization) and parameterizations.

What can we learn from this exercise? In order to understand business cycle fluctuations in developed and developing countries, it is crucial to construct theoretical models where primitive shocks generate fluctuations in efficiency and labor wedges. This implies that researchers need to build models where explicit frictions can be mapped in terms of the efficiency wedge A_t in (II.4) and/or the labor wedge $1 - \tau_{n,t}$ in (II.6). Of course, the models with these class of frictions also need to be able to explain the differences between developed and developing countries at business cycle frequencies. The following section discusses two widely used models to understand business cycle facts in developed and developing economies: the Aguiar and Gopinath (2007) model, and the Neumeyer and Perri (2005) model. As will become clear below, these models include frictions that can be interpreted in terms of efficiency and labor wedges in the prototype economy.

[16] For an open economy version of the prototype economy, Lama (2011) finds that neither the investment nor the bond wedge are able to explain output drops in Latin American countries.

[17] See, e.g., Chari et al. (2007) for the U.S.; Kersting (2008) for the U.K.; Lama (2011) for a set of Latin American countries.

II.4 What Explains the Differences at Business Cycle Frequencies?

The explanations behind understanding business cycles in developed and developing countries can be divided into two basic approaches. The first approach is identified with the work of Aguiar and Gopinath (2007). The basic idea is that business cycles may be explained within a neoclassical framework with no explicit policy and/or market distortions, where macroeconomic fluctuations are solely driven by shocks to total factor productivity (TFP). This specification does not imply that market imperfections are unimportant, but that such frictions may be well captured by a shock to TFP. The second approach is based on the work by Neumeyer and Perri (2005), Uribe and Yue (2006) and García-Cicco et al. (2010). These authors argue that models with explicit frictions other than TFP are well suited to explain the stylized facts in both types of economies.

II.4.1 The Aguiar and Gopinath (2007) model

Aguiar and Gopinath (A&G, 2007) derive a small open economy model with a single good, which is augmented to include transitory and permanent productivity shocks. Specifically, technology is described by a Cobb-Douglas production function of the form:

$$y_t = e^{z_t} k_t^{1-\alpha} \left(\Gamma_t l_t\right)^{\alpha}, \tag{II.10}$$

where $\alpha \in (0, 1)$ is the labor share of per capita output y_t, and capital and labor are denoted k_t and l_t, respectively. The productivity process is captured by the variables z_t and Γ_t. The term z_t represents the transitory component of productivity, which is given by the following first-order autoregressive process:

$$z_t = \rho_z z_{t-1} + \varepsilon_t^z, \tag{II.11}$$

with $|\rho_z| < 1$. Here, ε_t^z is an i.i.d. (transitory) technology shock distributed normally with zero mean and standard deviation σ_z. The permanent component of productivity is represented by Γ_t, which follows the law of motion:

$$\Gamma_t = e^{\psi_t} \Gamma_{t-1} = \prod_{s=0}^{t} e^{\psi_s}. \tag{II.12}$$

The term ψ_t is subject to a stochastic first-order autoregressive process denoted by

$$\psi_t = \left(1 - \rho_\psi\right) \mu_\psi + \rho_\psi \psi_{t-1} + \varepsilon_t^\psi, \tag{II.13}$$

with $\left|\rho_{\psi}\right| < 1$ and ε_t^{ψ} is an i.i.d. (permanent) productivity shock drawn from a normal distribution with zero mean and standard deviation σ_{ψ}. The expression μ_{ψ} is the productivity's long-run mean growth rate. The terms ε_t^z and ε_t^{ψ} constitute the two sources of fluctuations in the model.

There is also a representative household that lives forever, and derives expected utility from consumption and leisure given by a Cobb-Douglas utility function. Output produced can be allocated to consumption, investment, and internationally traded bonds. Capital depreciates at some constant rate, and investment is subject to a quadratic adjustment cost function.[18] When comparing expressions (II.4) and (II.10), it can be shown that the efficiency wedge A_t is equivalent to:

$$A_t \equiv \exp\left(z_t + \alpha\Psi_t\right), \tag{II.14}$$

where $\Psi_t \equiv \sum_{s=0}^{t}\left(\rho_{\psi}\psi_{s-1} + \varepsilon_s^{\psi}\right)$. In terms of the A&G model, the efficiency wedge takes the form of productivity disturbances captured by z_t and Ψ_t.

The next issue is to examine to what extent the model can reasonably explain some particular moments in the data, while at the same time accounting for business cycle differences between countries. A&G calibrate their model using data for Mexico and Canada, which are taken as representative cases of a developing and developed economy. Their results are presented in Table II.4. In general, the model does a good job at replicating the moments in the data for each case. In addition, the model explains some of the major differences already identified in Section II.2 above: in developing countries, output is relatively more volatile, consumption is more volatile than output, and net exports are strongly countercyclical.

In terms of the model, differences in business cycles between developed and developing countries are essentially explained by differences in the volatility of permanent productivity shocks σ_{ψ}.[19] In particular, A&G estimate this parameter using data for each country. For Mexico, they estimate a value for σ_{ψ} of 2.13 percent (with a standard error of 0.29); and for Canada, the estimate is 0.47 percent (with a standard error of 0.37). Thus, a more volatile shock to the trend component of productivity translates into a more volatile efficiency wedge via Ψ (see expression (II.14)). The intuition behind this result follows from the permanent income hypothesis. Suppose that an economy experiences a positive trend shock. This implies a temporary increase in the economy's growth rate, given

[18] Adjustment costs are commonly used in the small open-economy business cycle literature in order to avoid excessive volatility of investment. For further details see Mendoza (1991).

[19] More precisely, the relative importance of the random walk component of the Solow residual is larger in developing economies. See Aguiar and Gopinath (2007) for details.

Table II.4. Moments in the Aguiar and Gopinath (2007) model

	Emerging: Mexico		*Developed: Canada*	
	Data	*Model*	*Data*	*Model*
$\sigma(y)$	2.40	2.13	1.55	1.24
$\sigma(\Delta y)$	1.52	1.42	0.80	0.82
$\sigma(c)/\sigma(y)$	1.26	1.10	0.74	0.76
$\sigma(I)/\sigma(y)$	4.15	3.83	2.67	3.14
$\sigma(NX)/\sigma(y)$	0.90	0.95	0.57	0.65
$\rho(y)$	0.83	0.82	0.93	0.81
$\rho(\Delta y)$	0.27	0.18	0.55	0.17
$\rho(y, NX)$	−0.75	−0.50	−0.12	−0.15
$\rho(y, c)$	0.92	0.91	0.87	0.87
$\rho(y, I)$	0.91	0.80	0.74	0.82

Notes: Standard deviation of variable x, $\sigma(x)$, is reported in percentage terms. $\rho(x, z)$ is the correlation between variables x and z. y is output; Δy is output growth; c is consumption; I is investment; and NX are net exports over GDP.

Source: Aguiar and Gopinath (2007).

that $\rho_\psi \in (0, 1)$, as is actually the case from A&G's empirical estimates. The economy thus experiences a boost to current output, and an even larger boost to future output. As households expect a period of higher growth, consumption will respond more than income. This effect reduces savings and generates a large trade deficit. If the shock is instead transitory, households will increase consumption only slightly. Consequently, savings increase and the trade balance deteriorates by a smaller amount. According to A&G, the premise behind this result is that developing economies are characterized by frequent regime switches (in terms of reversals in fiscal, monetary, trade policies, etc...), in sharp contrast with developed countries. As a result, trend shocks are the major source of fluctuations in these countries.

Recently, Boz et al. (2011) modify the A&G model to incorporate imperfect information and learning about the underlying shocks in the economy. In particular, they take the same specification for trend and transitory shocks given by expressions (II.11)-(II.13). However, agents are now imperfectly informed about the true decomposition of productivity shocks into their transitory and permanent components. They observe a noisy signal on the trend shock, where the signal represents the degree of information imperfection. The way agents solve this signal extraction problem is by using the Kalman filter.[20] Therefore, anytime there is a signal in the productivity growth rate agents update their beliefs about the transitory component of productivity (i.e., agents *learn*). Using data for Mexico and Canada, structural estimates suggest that the degree of information imperfection is higher in developing economies. Furthermore, if the noisiness (variance) of the trend growth signals decreases in the baseline imperfect information model for Mexico, the relative variability of consumption falls and the trade balance is less countercyclical. In contrast to A&G, the model of Boz et al. (2011) does not need to rely on a high volatility in the trend shock to match the data. This suggests that the degree of uncertainty that agents face at the time of formulating expectations can help explain business cycle differences between developing and developed countries.

In a related paper, García-Cicco et al. (2010) test the empirical validity of a (slightly modified) version of the A&G model by using long-time series for developing countries. Specifically, the authors employ Greenwood-Hercowitz-Huffman (GHH, 1988) preferences, and estimate the parameters associated to the stochastic process of productivity using data for Argentina and Mexico over the period 1900–2005. The motivation for using a data span of more than 100 years is that such parameters can be more accurately estimated, especially those related to the permanent productivity shock. As in A&G, the stochastic productivity processes are given by (II.11) and (II.13). The parameters are estimated using Bayesian techniques for Argentina and GMM techniques for Mexico. However, García-Cicco et al. (2010) do not HP filter the data. Instead, their estimation procedure uses growth rates of output, consumption, and investment, as well as the trade balance-to-output ratio. When estimated over the long sample, they find that the modified version of the A&G model does a poor job at explaining business cycles in developing countries.

In trying to understand why the model does a good job at replicating short-sample moments, but a poor job at matching long-sample moments, the authors take the same dataset as used by A&G. The parameters are estimated using GMM techniques and the data is not HP filtered. They find that the model replicates the data quite well, thereby suggesting that differences in econometric techniques cannot account for the poor job of the model at matching the data. Arguably, the advantage of using a longer time span is that business cycles may be characterized more precisely.[21] For the period 1980–2003, both consumption and the

[20] See Ljungqvist and Sargent (2004) for an introduction to the Kalman filter.

[21] For most Latin American economies, García-Cicco et al. (2010) report that the period 1980–2003 contains only one and a half to two business cycles. In contrast, the period 1900–2005 has about nine business cycles. The authors also argue that the amplitude and frequency of business cycles in developing economies has been stable over the past century.

trade balance-to-output ratio are 1.25 and 1.15 times more volatile than output in Mexican data. This is not observed in the longer period 1900–2005: consumption is 1.5 times more volatile than output but the trade balance-to-output ratio is actually less volatile than output. As shown by García-Cicco et al. (2010), the A&G model has the property that the volatility of both consumption and the trade balance-to-output ratio go hand in hand if the persistence and/or volatility of the trend productivity shock is increased. Therefore, it is easier for the model to account for the data in the 1980–2003 period but not for the whole sample.

II.4.2 The Neumeyer and Perri (2005) model

Neumeyer and Perri (NP, 2005) derive a standard neoclassical small open economy model with a single good, where the only asset traded in international financial markets is a non-contingent bond. This bond is traded by both domestic firms and households. Firms trade in this asset as they need to pay the wage bill in advance of output production (i.e., they require working capital). This type of friction only applies to labor services; firms make payments to the owners of capital at the end of period when production is realized. There are also two shocks in the model: a TFP shock and an interest rate shock.

At the beginning of period t, firms hire labor l_t and capital k_t to produce output y_t. Workers earn w_t goods per unit of time so the wage bill is $w_t l_t$. Firms need to borrow working capital due to a technology friction that impedes transferring resources to the households that provide labor services. In particular, firms need to set aside a fraction θ of the wage bill at the beginning of the period, and a fraction $1-\theta$ at the end. Thus, firms have to borrow $\theta w_t l_t$ units of goods at the gross interest rate R_{t-1}. At the end of time t, firms pay $(1-\theta)\, w_t l_t$ out of labor services, $r_t k_t$ out of rental services to the owners of capital, and $\theta w_t l_t R_{t-1}$ out of repayments due to the working capital loan plus interest. Firm's profits π_t are thus given by:

$$\pi_t = y_t - w_t l_t - r_t k_t - \left(R_{t-1} - 1\right) \theta w_t l_t, \qquad \text{(II.15)}$$

where $\left(R_{t-1} - 1\right) \theta w_t l_t$ is the net interest on the fraction of the wage bill that must be paid with borrowed funds. In the model, a higher interest rate R raises the cost of labor and thus induces a fall in labor demand. In expression (II.15), the production function is Cobb-Douglas:

$$y_t = e^{z_t} F\left(k_t, (1+\gamma)^t l_t\right) = e^{z_t} k_t^{1-\alpha} \left[(1+\gamma)^t\, l_t\right]^{\alpha}, \qquad \text{(II.16)}$$

where γ is the deterministic growth rate of labor-augmenting technological change and e^{z_t} is the stochastic component of productivity.

Households maximize lifetime expected utility from consumption and leisure. They also derive income from labor and rental services, and interest payments on the internationally traded bond. The interest rate paid on bonds at time t is just R_{t-1}. Total income may be allocated to consumption, investment, and bonds. Capital depreciates at some constant rate, and it is subject to adjustment costs. The model is closed by assuming a convex cost of holding bonds.

The interest rate R_t at which foreigners are willing to lend resources to domestic agents has two sources of volatility: the preference of international investors for risky assets, and the default risk on payments from domestic agents to foreigners. Accordingly, the interest rate may be decomposed as follows:

$$R_t = R_t^* D_t. \tag{II.17}$$

In expression (II.17), R^* is the international rate paid on risky assets (not specific to a particular developing country), and D is the country spread over R^* paid by borrowers in a particular economy.[22] The idea behind the country spread is that loans to the domestic economy are risky assets, as there is a probability that agents default on their obligations. The higher the probability of default, the higher the country spread. Productivity disturbances follow an AR(1) process given by equation (II.11). A similar process for the international interest rate R^* is assumed. However, the country spread D has two alternative specifications: a standard AR(1) process (the *independent country risk* scenario), and an ad-hoc function of expected productivity in $t + 1$ (the *induced country risk* scenario). In particular, the authors set $D_t = \eta\left(\mathrm{E}_t e^{z_{t+1}}\right)$. If $\widehat{x}$ represents the percentage deviation of variable x from its balanced-growth path, a log-linear form of this specification may be written as

$$\widehat{D}_t = \overline{\eta}\mathrm{E}_t\left(\widehat{z}_{t+1}\right) + \varepsilon_t^d, \tag{II.18}$$

where $\overline{\eta} < 0$ is a parameter that captures how much country risk responds to expected productivity shocks and ε_t^d is a country risk i.i.d. shock distributed with zero mean and constant variance. The intuition for $\overline{\eta} < 0$ is simple: the expectation of a positive productivity shock decreases the probability of default and thus the risk premia D.

[22] Uribe and Yue (2006) adopt an alternative approach to estimate the stochastic process for R. They estimate a first-order VAR system for output, investment, the trade balance-to-output ratio, the gross real US interest rate, and the gross real (emerging) country interest rate. This method defines a law of motion for R in terms of contemporary and one-period lag variables included in the VAR system. They do not introduce technology shocks into their specification.

Table II.5. Selected moments in the Neumeyer and Perri (2005) model: Argentina

	Data	*Independent country risk*	*Induced country risk*
$\sigma(TC)/\sigma(GDP)$	1.17	1.13	1.54
$\rho(GDP, RIR)$	−0.63	−0.29	−0.54
$\rho(GDP, NX)$	−0.89	−0.08	−0.80
$\rho(GDP, TC)$	0.97	0.87	0.97
$\rho(GDP, INV)$	0.94	0.44	0.90
$\rho(RIR, NX)$	0.71	0.96	0.65
$\rho(RIR, TC)$	−0.67	−0.70	−0.60

Notes: $\sigma(x)$ is the standard deviation of variable x. $\rho(x, z)$ is the correlation between variables x and z. See the notes in Table II.2 for a description of the variables.

Source: Neumeyer and Perri (2005).

To understand how the N&P model has the potential to account for business cycles, it is useful to look at the marginal rate of substitution between leisure and consumption:

$$-\frac{U_{l,t}}{U_{c,t}} = w_t = \frac{e^{z_t}(1+\gamma)^t F_l\left(k_t, l_t\right)}{1+\theta\left(R_{t-1}-1\right)}, \tag{II.19}$$

where $F_l\left(k_t, l_t\right)$ is the marginal product of labor. After comparing equations (II.6) and (II.19) it may be shown that the labor wedge is equivalent to:

$$1-\tau_{n,t} \equiv \frac{1}{1+\theta(R_{t-1}-1)}. \tag{II.20}$$

Hence, fluctuations in the labor wedge are derived from fluctuations in the interest rate R. In addition, fluctuations in the efficiency wedge are explained by standard fluctuations in productivity e^{z_t}. The combination of these two wedges is what makes this model a good candidate to explain business cycle facts in developed and developing countries.[23]

[23] The model also requires the elimination of the wealth effect on labor supply so that employment and interest rates are negatively correlated. See Neumeyer and Perri (2005) for details.

Table II.5 presents the results of the N&P model using data for Argentina during the period 1983–2001. In general, the induced country risk model does a better job than the independent country risk model at replicating the data.[24] For example, the former yields correlations of output with the interest rate, net exports, and investment more consistent with the data. This finding suggests that country risk is induced by domestic fundamentals. At the same time, country risk amplifies the effects of fundamental shocks on business cycles through the working capital mechanism. As N&P acknowledge, the way expected productivity affects the country risk premia is a type of "reduced form"approach to endogenous default. However, their results imply that in order to understand business cycles in developing countries, it is crucial to have a more elaborated model to disentangle the mechanism through which shocks to fundamentals affect country risk spreads.[25]

Recently, Chang and Fernández (2010) take the N&P model and extend it to include transitory and permanent productivity shocks in the spirit of the A&G model. In particular, this *encompassing* model replaces the production function (II.16) with the A&G function (II.10) so that trend shocks are incorporated into the analysis. They also replace the induced country risk specification (II.18) by the following expression:

$$\widehat{D}_t = \eta_1 \mathrm{E}_t \widehat{z}_{t+1} + \eta_2 \mathrm{E}_t \widehat{\varphi}_{t+1} + \varepsilon_t^d, \qquad \text{(II.21)}$$

with parameters η_1 and η_2 taking negative values, and $\widehat{\varphi}_{t+1}$ following the first-order autoregressive process given by (II.13). Notice that the stochastic process (II.18) is now a particular case of (II.21) with $\eta_1 \equiv \overline{\eta}$ and $\eta_2 = 0$. To the extent that $\eta_2 < 0$, Chang and Fernández (2010) allow the possibility that trend shocks have a negative impact on country risk. This is an encompassing model in the sense that it combines stochastic trends (as in A&G) with interest rate shocks in a working capital environment (as in N&P). Therefore, the model can then be used to evaluate the relative contribution of interest rate, trend, and transitory productivity shocks to explain the data.

Relevant parameters, including those for the exogenous shock processes, are estimated by Bayesian methods, and the authors use the same dataset as A&G for Mexico for comparability purposes. The results of Chang and Fernández (2010)

[24] A similar result is reported in Antón and Villegas (2013) using data for Mexico.

[25] García-Cicco et al. (2010) also report that the country risk premium is important to account for fluctuations in developing countries in a model with no working capital. Their model also features a "reduced form" approach to country risk shocks. Mendoza and Yue (2012) feature a sovereign default setting with working capital where country risk endogenously amplifies the effects of adverse productivity shocks on output, and vice-versa. As the sovereign default decision is explicitly taken into account, the model provides an explicit mechanism to understand these intricate relationships.

suggest that transitory productivity shocks account for most of the macroeconomic fluctuations in developing economies. However, interest rate shocks also play an important role as well, explaining between 6 and 10 percent of the variance in output and consumption, one fourth of the variance of investment, and about half of the variance in the trade balance-to-output ratio. In sharp contrast, trend shocks only explain three percent or less of the variance of such variables. As shown by the authors, these results are robust to a series of alternative specifications. Overall, these findings support those of N&P in the sense that explaining fluctuations in developing economies requires financial frictions in the form of working capital that amplify transitory technology shocks. In terms of the benchmark prototype model of Chari et al. (2007), this means that explaining fluctuations in developing countries requires a combination of efficiency and labor wedges.

II.5 Conclusions

This chapter has reported a number of stylized facts at business cycle frequencies for both developed and developing economies. It has also presented a prototype economy to interpret macroeconomic fluctuations based on a series of distortions, or wedges, to an otherwise standard neoclassical growth model. Finally, two well-known models used to understand business cycle facts in developed and developing economies have been discussed through the lens of the prototype economy of Chari et al. (2007).

However, there are some economic phenomena that are also important to understand fluctuations in developing economies that have not been discussed here. One of them is the role of sharp contractions of international capital flows (i.e., sudden stops) as discussed by Calvo (1998) and Calvo et al. (2004). These events are associated with large depreciations and financial disruptions in small open economies leading to large output falls. In the models of Christiano et al. (2004) and Mendoza and Smith (2006), agents face an external credit constraint; when such a constraint suddenly binds, domestic interest rates increase, output falls, and there is a large increase in the current account surplus (the sudden stop in capital flows). The second phenomena is related to sovereign defaults in a context of a large external debt as a share of output. As documented by Mendoza and Yue (2012), default events are associated with deep recessions in developing economies. For example, the models of Aguiar and Gopinath (2006), Arellano (2008), Yue (2010), and Cuadra et al. (2010), feature a sovereign borrower facing output shocks. Debt contracts are not enforceable since the borrower always has the option to default on them. If the borrower optimally chooses to default on its debt (due, e.g., to a negative shock), the economy experiences an output loss. Even though these events particular to developing countries are not discussed in detail here, hopefully the framework analyzed in this chapter may give the reader a good starting point to understand such phenomena.

Contrasting business cycles facts in developing and developing countries is a very active area of research. Based on the evidence provided in this chapter, theoretical models including distortions in the production function and the marginal rate of substitution between leisure and consumption are the most promising mechanisms to account for macroeconomic fluctuations in developed and developing economies.

REFERENCES

Aguiar, M., Gopinath, G. (2007). Emerging Market Business Cycles: The Cycle is the Trend. *Journal of Political Economy* 115, 69–102.

Aguiar, M., Gopinath, G. (2006). Defaultable Debt, Interest Rates and the Current Account. *Journal of International Economics* 69, 64–83.

Aiolfi, M., Catao, L. A. V., Timmermann, A. (2011). Common Factors in Latin America's Business Cycles. *Journal of Development Economics* 95, 212-228.

Antón, A. (2011). El Ciclo Económico en México: Características y Perspectivas. *Realidad, Datos y Espacio* 2, 32–49.

Antón, A. (2008). Accounting for Output Fluctuations in Mexico. Working Paper: 2008-05, Banco de México.

Antón, A., Villegas, A. (2013). El Papel de la Tasa de Interés Real en el Ciclo Económico de México. *El Trimestre Económico* 80, 773–803.

Arellano, C. (2008). Default Risk and Income Fluctuations in Emerging Economies. *American Economic Review* 98, 690–712.

Artis, M. J., Kontolemis, Z. G., Osborn, D. R. (1997). Business Cycles for G7 and European Countries. *Journal of Business* 70, 249–279.

Backus, D. K., Kehoe, P. J. (1992). International Evidence on the Historical Properties of Business Cycles. *American Economic Review* 82, 864–888.

Backus, D., Kehoe, P. J., Kydland, F. E. (1995). International Business Cycles: Theory and Evidence. In Cooley, T. F., ed., *Frontiers in Business Cycle Research*. Princeton University Press.

Baxter, M., King, R. G. (1999). Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *Review of Economics and Statistics* 81, 575–593.

Boz, E., Daude, C., Durdu, C. B. (2011). Emerging Market Business Cycles: Learning about the Trend. *Journal of Monetary Economics* 58, 616–631.

Bry, G., Boschan, C. (1971). *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. New York: National Bureau of Economic Research.

Burns, A. F., Mitchell, W. C. (1946). *Measuring Business Cycles*. New York: National Bureau of Economic Research.

Calderón, C., Fuentes, R. (2010). Characterizing the Business Cycles of Emerging Countries. World Bank Policy Research Working Paper No. 5343. World Bank.

Calvo, G. A. (1998). Capital Flows and Capital-Market Crises: The Simple Economics of Sudden Stops. *Journal of Applied Economics* 1, 35–54.

Calvo, G. A., Izquierdo, A., Mejía, L. F. (2004). On the Empirics of Sudden Stops: The Relevance of Balance Sheet Effects. NBER Working Paper No. 10520. National Bureau of Economic Research.

Canova, F. (2007). *Methods for Applied Macroeconomic Research*. Princeton, NJ: Princeton University Press.

Chang, R., Fernández, A. (2010). On the Sources of Aggregate Fluctuations in Emerging Economies. NBER Working Paper No. 15938. National Bureau of Economic Research.

Chari, V. V., Kehoe, P. J., McGrattan, E. (2007). Business Cycle Accounting. *Econometrica* 75, 781–836.

Chari, V. V., Kehoe, P. J., McGrattan, E. (2006). Appendices: Business Cycle Accounting. Staff Report No. 362. Federal Reserve Bank of Minneapolis.

Christiano, L. J., Davis, J. M. (2006). Two Flaws in Business Cycle Accounting. NBER Working Paper No. 12647. National Bureau of Economic Research.

Christiano, L. J., Gust, C., Roldos, J. (2004). Monetary Policy in a Financial Crisis. *Journal of Economic Theory* 119, 64–103.

Cuadra, G., Sanchez, J. M., Sapriza, H. (2010). Fiscal Policy and Default Risk in Emerging Markets. *Review of Economic Dynamics* 13, 452–469.

García-Cicco, J., Pancrazi, R., Uribe, M. (2010). Real Business Cycles in Emerging Countries. *American Economic Review* 100, 2510–2531.

Greenwood, J., Hercowitz, Z., Huffman, G. W. (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review* 78, 402–417.

Harding, D., Pagan, A. (2002). Dissecting the Cycle: A Methodological Investigation. *Journal of Monetary Economics* 29, 365–381.

Hodrick, R. J., Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29, 1–16.

Kersting, E. K. (2008). The 1980s Recession in the UK: A Business Cycle Accounting Perspective. *Review of Economic Dynamics* 11, 179–191.

King, R. G., Plosser, C. I., Rebelo, S. T. (1988). Production, Growth and Business Cycles I. The Basic Neoclassical Model. *Journal of Monetary Economics* 21, 195–232.

Kydland, F., Prescott, E. C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica* 50, 1345–1370.

Lama, R. (2011). Accounting for Output Drops in Latin America. *Review of Economic Dynamics* 14, 295–316.

Ljungqvist, L., Sargent, T. J. (2004). *Recursive Macroeconomic Theory*, The MIT Press, 2nd. edition.

Lucas, R. E. (1977). Understanding Business Cycles. *Carnegie-Rochester Conference Series on Public Policy* 5, 7–29.

Lucas, R. E. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.

Mendoza, E. G. (1995). The Terms of Trade, the Real Exchange Rate, and Economic Fluctuations. *International Economic Review* 36, 101–137.

Mendoza, E. G. (1991). Real Business Cycles in a Small Open Economy. *American Economic Review* 81, 797–818.

Mendoza, E. G., Yue, V. Z. (2012). A General Equilibrium Model of Sovereign Default and Business Cycles. *Quarterly Journal of Economics* 127, 889–946.

Mendoza, E. G., Smith, K. A. (2006). Quantitative Implications of a Debt-Deflation Theory of Sudden Stops and Asset Prices. *Journal of International Economics* 70, 82–114.

Neumeyer, P. A., Perri, F. (2005). Business Cycles in Emerging Economies: The Role of Interest Rates. *Journal of Monetary Economics* 52, 345–380.

Plosser, C. I. (1989). Understanding Real Business Cycles. *Journal of Economic Perspectives* 3, 51–77.

Prescott, E. C. (1986). Theory Ahead of Business-Cycle Measurement. *Carnegie-Rochester Conference Series on Public Policy* 25, 11–44.

Rand, J., Tarp, F. (2002). Business Cycles in Developing Countries: Are They Different? *World Development* 30, 2071–2088.

Talvi, E., Végh, C. A. (2005). Tax Base Variability and Procyclical Fiscal Policy in Developing Countries. *Journal of Development Economics* 78, 156–190.

Uribe, M., Yue, V. (2006). Country Spreads and Emerging Countries: Who Drives Whom? *Journal of International Economics* 69, 6–36.

Yue, V. Z. (2010). Sovereign Default and Debt Renegotiation. *Journal of International Economics* 80, 176–187.

III. DESIGNING MONETARY POLICY RULES FOR SMALL OPEN ECONOMIES

STEPHEN MCKNIGHT
El Colegio de México

III.1 INTRODUCTION

Since Taylor (1993), it is now commonplace to think of monetary policy in terms of a feedback rule that links adjustments in the nominal interest rate to movements in inflation. The popularity of conducting monetary policy in this way is such that inflation targeting feedback rules are now the defining characteristic of monetary policymaking worldwide (Huang et al., 2009). Consequently, a key issue in monetary economics relates to the design of these feedback rules. Of particular importance is that the monetary policy rule adopted by the central bank should do no harm; that is the feedback rule should avoid generating multiple equilibria or equilibrium indeterminacy, which can destabilize the economy through the emergence of expectations-driven fluctuations:

> If one evaluates policy rules according to how bad is the worst outcome that they might allow, it would be appropriate to assign an absolute priority to the selection of a rule that would guarantee determinacy of equilibrium. [Woodford, 2003, p. 89]

It has been well established in the closed economy literature that by adhering to the *Taylor principle*, i.e. a policy that adjusts the nominal interest rate by proportionally more than the increase in inflation, a central bank can avoid indeterminacy of equilibrium.[1] Recently, a number of studies have investigated whether the Taylor principle guarantees a determinate rational expectations equilibrium in open economies.[2] As pointed out by Llosa and Tuesta (2008) there are two additional issues that central banks must consider when designing monetary policy rules for open economies. The first issue relates to the operational choice of the

[1] See, e.g., Bernanke and Woodford (1997), Clarida et al. (2000), Benhabib et al. (2001), Carlstrom and Fuerst (2001, 2005), Woodford (2003), Lubik and Marzo (2007).

[2] See, e.g., Batini et al. (2004), De Fiore and Liu (2005), Linnemann and Schabert (2006), Llosa and Tuesta (2008), Bullard and Scaling (2009), Leith and Wren-Lewis (2009), McKnight (2011a, 2011b), McKnight and Mihailov (2012).

inflation indicator used by the central bank in the policy rule. Specifically, should central banks use consumer price inflation or producer price inflation when setting interest rates?[3,4] The second issue relates to whether the policy rule should react directly to changes in either the nominal or real exchange rate. The purpose of this chapter is to summarize the key findings and conclusions of this literature with the aim of providing a reference for those policymakers, researchers, and students interested in the issue of equilibrium (in)determinacy. There are two reasons why the focus of attention rests solely with *small* open economies. The first is for simplicity. The determinacy conditions of small open economy models are significantly easier to analytically derive than their multi-country counterparts. Secondly, the small open economy assumption is appropriate for most developed and developing countries, including Mexico.

The next section outlines the baseline model which will be used to conduct the analysis of monetary policy design. It is a small open economy version of the popular New Keynesian modeling framework; that is, a dynamic stochastic general equilibrium model with imperfect competition and Calvo-style price-stickiness. The openness of the economy to international trade is proxied by the degree of home bias in consumption. While the law of one price holds continuously, home bias generates deviations in purchasing power parity, thereby permitting fluctuations in the terms of trade and the real exchange rate. The baseline model follows the vast majority of the existing literature and assumes a cashless economy where labor is the only factor of production. In Section III.3 the determinacy analysis is conducted for the baseline model under alternative specifications of the monetary policy rule. This section first illustrates the key insight of De Fiore and Liu (2005), i.e. that economies open to international trade are more susceptible to indeterminacy than closed economies. It then discusses the policy rule specifications that generate this important result, and how the feedback rule can be designed to minimize the indeterminacy problem. Two main policy recommendations emerge from this analysis. First, managed exchange rate feedback rules, by helping to alleviate indeterminacy, can have a stabilizing effect on the economy. Second, central banks should adhere to policy rules that adjust the nominal interest rate to movements in producer price inflation rather than consumer price inflation. In Section III.4 the analysis is extended by introducing money, and adding capital and investment spending to the baseline model. It is shown that the previous policy recommendations are highly sensitive to these modeling assumptions, such that now feedback rules that react to producer price inflation

[3] In closed economies consumer price inflation and producer price inflation are the same concept.

[4] In the literature producer price inflation is sometimes referred to as domestic price inflation.

can potentially be more destabilizing than policy rules that react to consumer price inflation. Finally, Section III.5 briefly concludes.

III.2 A SMALL OPEN ECONOMY MODEL

This section outlines the baseline small open economy model, which is based on Gali and Monacelli (2005) and Faia and Monacelli (2008). Consider a global economy that is populated by identical, infinitely-lived households of measure one. All households reside in two countries: a small open economy, denoted home, and the rest of the world, denoted foreign. The home country has a population of measure n, and the foreign country of measure $1-n$. Preferences and technologies are assumed to be symmetric across the two countries. In what follows asterisks denote foreign variables, and subscripts H (F) denote variables of home (foreign) origin.

III.2.1 Households

Consumption indexes. In the home economy households consume a composite of domestic C_H and imported C_F consumption bundles of goods composed of imperfectly substitutable varieties:

$$C_t \equiv \left[(1-\gamma)^{\frac{1}{\theta}} C_{H,t}^{\frac{\theta-1}{\theta}} + \gamma^{\frac{1}{\theta}} C_{F,t}^{\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}; \tag{III.1}$$

$$C_{H,t} \equiv \left[\left(\frac{1}{n}\right)^{\frac{1}{\lambda}} \int_0^n C_{H,t}(i)^{\frac{\lambda-1}{\lambda}} di\right]^{\frac{\lambda}{\lambda-1}};$$

$$C_{F,t} \equiv \left[\left(\frac{1}{1-n}\right)^{\frac{1}{\lambda}} \int_n^1 C_{F,t}(i)^{\frac{\lambda-1}{\lambda}} di\right]^{\frac{\lambda}{\lambda-1}}.$$

The parameter $\theta > 0$ measures the elasticity of substitution between home and foreign goods. The parameter $\gamma \equiv (1-n)a$ captures the home country's preferences for imported goods, which depends on the relative size of the foreign country $(1-n)$ and the degree of trade openness $a \in (0, 1)$. The parameter $\lambda > 1$ measures the elasticity of substitution between the varieties of goods produced within

H or F, where $i \in [0, 1]$. In an analogous manner, the aggregate consumption basket in the foreign country is given by:

$$C_t^* \equiv \left[(1-\gamma^*)^{\frac{1}{\theta}} C_{F,t}^{*\frac{\theta-1}{\theta}} + \gamma^{*\frac{1}{\theta}} C_{H,t}^{*\frac{\theta-1}{\theta}}\right]^{\frac{\theta}{\theta-1}}, \tag{III.2}$$

where $\gamma^* \equiv na$. Since $a < 1$, this implies home bias in consumption i.e. $1-\gamma > \gamma^*$.

The optimal allocation of expenditures between domestic and imported goods yields the following aggregate demand conditions for the two countries:

$$C_{H,t} = (1-\gamma)\left(\frac{P_{H,t}}{P_t}\right)^{-\theta} C_t; \qquad C_{F,t} = \gamma\left(\frac{P_{F,t}}{P_t}\right)^{-\theta} C_t; \tag{III.3}$$

$$C_{F,t}^* = (1-\gamma^*)\left(\frac{P_{F,t}^*}{P_t^*}\right)^{-\theta} C_t^*; \qquad C_{H,t}^* = \gamma^*\left(\frac{P_{H,t}^*}{P_t^*}\right)^{-\theta} C_t^*, \tag{III.4}$$

where the demand for individual goods i is given by:

$$C_{H,t}(i) = \frac{1}{n}\left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\lambda} C_{H,t}; \qquad C_{F,t}(i) = \frac{1}{1-n}\left(\frac{P_{F,t}(i)}{P_{F,t}}\right)^{-\lambda} C_{F,t}; \tag{III.5}$$

$$C_{F,t}^*(i) = \frac{1}{1-n}\left(\frac{P_{F,t}^*(i)}{P_{F,t}^*}\right)^{-\lambda} C_{F,t}^*; \qquad C_{H,t}^*(i) = \frac{1}{n}\left(\frac{P_{H,t}^*(i)}{P_{H,t}^*}\right)^{-\lambda} C_{H,t}^*. \tag{III.6}$$

Price indexes, the terms of trade, and the real exchange rate. Given the consumption indexes (III.1) and (III.2), the corresponding home P_t and foreign P_t^* consumer price indexes can be derived as:

$$P_t = \left[(1-\gamma)P_{H,t}^{1-\theta} + \gamma P_{F,t}^{1-\theta}\right]^{\frac{1}{1-\theta}}; \qquad P_t^* = \left[(1-\gamma^*)P_{F,t}^{*1-\theta} + \gamma^* P_{H,t}^{*1-\theta}\right]^{\frac{1}{1-\theta}}, \tag{III.7}$$

where P_H (P_H^*) is the price sub-index for home-produced consumption goods denominated in home (foreign) currency and P_F (P_F^*) is the price sub-index for foreign-produced consumption goods denominated in the home (foreign) currency. Assuming that there are no costs to trade, and the law of one price holds, implies

$$P_{H,t} = e_t P_{H,t}^*; \qquad P_{F,t} = e_t P_{F,t}^*, \tag{III.8}$$

where e_t denotes the nominal exchange rate. Using (III.8) and defining the terms of trade to be $T_t \equiv \frac{P_{F,t}}{P_{H,t}}$, the consumer price indexes given in (III.7) can be written as:

$$\frac{P_t}{P_{H,t}} = \left[(1-\gamma) + \gamma T_t^{1-\theta}\right]^{\frac{1}{1-\theta}}; \qquad \frac{P_t^*}{P_{F,t}^*} = \left[(1-\gamma^*) + \gamma^* T_t^{\theta-1}\right]^{\frac{1}{1-\theta}}, \tag{III.9}$$

and the real exchange rate can be expressed as:

$$Q_t \equiv \frac{e_t P_t^*}{P_t} = T_t \frac{P_t^*}{P_t}\frac{P_{H,t}}{P_{F,t}^*}. \tag{III.10}$$

Preferences and asset markets. In the home country, the representative household seeks to maximize expected lifetime utility:

$$\max \mathrm{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t U\left(C_t, L_t\right) \right\},$$

$$U\left(C_t, L_t\right) \equiv \frac{C_t^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \frac{L_t^{1+\omega}}{1+\omega},$$

where C_t is consumption, L_t is labor supply, E_0 is the conditional expectations operator, $0 < \beta < 1$ is the discount factor, $\sigma > 0$ is the intertemporal elasticity of substitution in consumption, and $\omega > 0$ is the inverse elasticity of labor supply. The household receives real income from wages w_t and nominal profits Π_t from the ownership of domestic firms. International asset markets are assumed to be complete.[5] The household can purchase a state-contingent nominal bond B_{t+1} that pays one unit of domestic currency in period $t+1$ if a specific state is realized at a period t price $\Gamma_{t,t+1}$. The period budget constraint is thus given by:

$$\mathrm{E}_t\{\Gamma_{t,t+1} B_{t+1}\} + P_t C_t \leq B_t + P_t w_t L_t + \Pi_t. \tag{III.11}$$

The first-order conditions from the representative household's maximization problem yield:

$$\beta \frac{U_c\left(C_{t+1}\right)}{U_c\left(C_t\right)} \frac{P_t}{P_{t+1}} = \Gamma_{t,t+1}, \tag{III.12}$$

[5] For analytical tractability we adopt the standard assumptions of the existing literature. Assuming incomplete international financial markets would not alter in any way the main results and conclusions of this chapter.

$$\frac{-U_l(L_t)}{U_c(C_t)} = w_t. \tag{III.13}$$

Equation (III.12) is the intertemporal optimality condition, which must hold for each possible state, and equation (III.13) is the optimal labor supply condition that equates the real wage to the marginal rate of substitution between consumption and labor. Optimizing behavior further implies that the budget constraint (III.11) holds with equality in each period, and the transversality condition (III.14) is satisfied:

$$\lim_{i\to\infty} \mathrm{E}_t\{\Gamma_{t,t+i} B_{t+i}\} = 0. \tag{III.14}$$

Letting R_t denote the gross nominal yield on a one-period riskless bond defined as $R_t^{-1} \equiv \mathrm{E}_t\{\Gamma_{t,t+1}\}$, equation (III.12) can be expressed as a conventional consumption Euler equation:

$$\beta \mathrm{E}_t \left\{ \frac{U_c(C_{t+1})}{P_{t+1}} \right\} = \frac{U_c(C_t)}{P_t}\frac{1}{R_t}. \tag{III.15}$$

The foreign representative household can also purchase home-currency denominated bonds (after currency conversion). Consequently, the foreign equivalent to (III.12) is:

$$\beta \frac{U_{c^*}(C^*_{t+1})}{U_{c^*}(C^*_t)} \frac{e_t P^*_t}{e_{t+1} P^*_{t+1}} = \Gamma_{t,t+1}. \tag{III.16}$$

Using (III.12) and (III.16), and iterating yields

$$Q_t = q_0 \frac{U_{c^*}(C^*_t)}{U_c(C_t)}, \tag{III.17}$$

where the constant $q_0 \equiv Q_0 \left[\frac{u_c(C_0)}{u_{c^*}(C^*_0)}\right]$. Equation (III.17), which follows from the assumption of complete asset markets, equates the marginal rate of substitution between home and foreign consumption to the relative price of the aggregate consumption baskets.

III.2.2 Firms

Production. Firms hire labor to produce output, given a real wage rate w_t. A firm of type i has a linear production technology:

$$Y_t(i) = L_t(i). \tag{III.18}$$

Given competitive prices of labor, cost minimization yields

$$mc_t = w_t \frac{P_t}{P_{H,t}}, \qquad \text{(III.19)}$$

where $mc_t \equiv \frac{MC_t}{P_{H,t}}$ is real marginal cost.

Price setting. Firms set prices according to Calvo (1983), where in each period there is a constant probability $1-\psi$ that a firm will be randomly selected to adjust its price, which is drawn independently of past history. A domestic firm i, faced with resetting its price at time t, chooses $\widetilde{P}_{H,t}(i)$ to maximize:

$$\max_{\widetilde{P}_{H,t}(i)} \mathrm{E}_t \sum_{s=0}^{\infty} (\beta\psi)^s X_{t,t+s} \left\{ \left[\widetilde{P}_{H,t}(i) - P_{H,t+s} mc_{t+s}\right] \left(\frac{\widetilde{P}_{H,t}(i)}{P_{H,t+s}}\right)^{-\lambda} \left(C_{H,t+s} + \frac{(1-n)}{n} C^*_{H,t+s}\right) \right\},$$

where the firm's stochastic discount factor used to value random date $t+s$ payoffs is $\beta^s X_{t,t+s} = [U_c(C_{t+s})/U_c(C_t)](P_t/P_{t+s})$. All firms that are given the opportunity to reset their price in period t, all behave in an identical manner. It follows that the optimal price setting decision is given by:

$$\widetilde{P}_{H,t} = \frac{\lambda}{\lambda-1} \frac{\mathrm{E}_t \sum_{s=0}^{\infty} (\beta\psi)^s X_{t,t+s} P_{H,t+s}^{\lambda+1} \left(C_{H,t+s} + \frac{(1-n)}{n} C^*_{H,t+s}\right) mc_{t+s}}{\mathrm{E}_t \sum_{s=0}^{\infty} (\beta\psi)^s X_{t,t+s} P_{H,t+s}^{\lambda} \left(C_{H,t+s} + \frac{(1-n)}{n} C^*_{H,t+s}\right)}, \qquad \text{(III.20)}$$

where the optimal price set is a mark-up $\frac{\lambda}{\lambda-1}$ over a weighted average of future marginal costs. The price sub-index evolves according to:

$$P_{H,t}^{1-\lambda} = \psi P_{H,t-1}^{1-\lambda} + (1-\psi)\widetilde{P}_{H,t}^{1-\lambda}. \qquad \text{(III.21)}$$

III.2.3 Central bank

Initially, it is assumed that the monetary authority adjusts the (gross) nominal interest rate in response to expected changes in consumer price inflation (CPI) according to the rule:

$$R_t = R \left(\frac{\mathrm{E}_t\{\pi_{t+1}\}}{\pi}\right)^{\mu}, \qquad \text{(III.22)}$$

where $\pi_{t+1} \equiv \frac{p_{t+1}}{p_t}$ denotes the CPI inflation rate in period $t+1$, $\mu \geq 0$ is the inflation response coefficient, and $R = \pi/\beta > 1$ and π denote the steady state nominal interest rate and steady state CPI inflation rate, respectively. The Taylor principle is represented by $\mu > 1$, implying that the nominal interest rate rises proportionally more than an increase in expected inflation. The justification for studying a monetary policy rule of this form is motivated by the following two stylized facts. First, the majority of small open economies define their monetary policy goals in terms of CPI inflation (see, e.g., De Fiore and Liu, 2005), and second, empirical evidence suggests that central banks conduct monetary policy in a forward-looking manner (see, e.g., Clarida et al., 1998, 2000; Orphanides, 2001, 2004; Mihailov, 2006).

III.2.4 Market clearing and equilibrium

Goods market clearing in the home country requires:

$$Y_t(i) = nC_{H,t}(i) + (1-n)C^*_{H,t}(i)$$

$$\Rightarrow Y_t(i) = \left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\lambda}\left(\frac{P_{H,t}}{P_t}\right)^{-\theta}\left[[1-(1-n)a]C_t + (1-n)aC^*_t Q^{\theta}_t\right], \quad \text{(III.23)}$$

after using equations (III.3)–(III.6), (III.8), and (III.10), along with the definitions of γ and γ^*. In order to characterize the small open economy, I take the limit of the home country size to zero i.e. $n \to 0$. Consequently, the equilibrium dynamics of the rest of the world are exogenous from the viewpoint of the home country. Inserting (III.23) into the definition of aggregate domestic output: $Y_t \equiv \left[\left(\frac{1}{n}\right)^{\frac{1}{\lambda}} \int_0^n Y_t(i)^{\frac{\lambda-1}{\lambda}} di\right]^{\frac{\lambda}{\lambda-1}}$ yields:

$$Y_t = \left(\frac{P_{H,t}}{P_t}\right)^{-\theta}\left[(1-a)C_t + aC^*_t Q^{\theta}_t\right], \quad \text{(III.24)}$$

where the CPI index for the foreign country given in (III.9) collapses to $P^*_t = P^*_{F,t}$. Consequently, the terms of trade identity (III.10) reduces to:

$$T_t = Q_t \frac{P_t}{P_{H,t}}. \quad \text{(III.25)}$$

Market clearing in the factor goods market requires $L_t = \int_0^n L_t(i)di$, and aggregating the production function across domestic firms yields:

$$Y_t = \frac{L_t}{d_t}, \quad \text{(III.26)}$$

where $d_t \equiv \int_0^n \left(P_{H,t}(i)/P_{H,t}\right)^{-\lambda} di$ measures the price dispersion of home-produced goods.

Equilibrium. Given an exogenous sequence for foreign consumption $\{C_t^*\}$, and an initial condition for d_{t_0-1}, an equilibrium for the small open economy consists of a sequence of prices $\{P_t, P_{H,t}, \widetilde{P}_{H,t}, d_t, w_t, mc_t, Q_t, T_t\}$, a sequence of allocations $\{C_t, Y_t, L_t\}$, and a monetary policy rule $\{R_t\}$ satisfying: (i) the optimality conditions of households (III.13) and (III.15), the international risk sharing condition (III.17), and the transversality condition (III.14); (ii) the optimality conditions of firms (III.19), (III.20), (III.21), (III.26), and a law of motion for price dispersion; (iii) the monetary policy rule (III.22); (iv) the goods market clears (III.24); (v) the consumer price index (III.9) and the terms of trade identity (III.25).

III.2.5 The steady state

For this small open economy, it can be shown that there exists a unique, zero-inflation, perfect-foresight steady state. In what follows all variables without time indices denote steady state values. In a steady state, the international risk sharing condition (III.17) and the terms of trade identity (III.25) imply:

$$C = Q^\sigma C^* = T^\sigma \left(\frac{P_H}{P}\right)^\sigma C^*, \qquad \text{(III.27)}$$

where I have normalized the constant $q_0 = 1$. Noting that $Y = L$ (since $d = 1$) and $mc = \frac{\lambda-1}{\lambda}$, labor market clearing implies from (III.13) and (III.19):

$$C^{\frac{1}{\sigma}} Y^\omega = w = mc\frac{P_H}{P} = \left(\frac{\lambda-1}{\lambda}\right)\frac{P_H}{P}. \qquad \text{(III.28)}$$

Using (III.27) to eliminate C from (III.28) gives:

$$Y = \left[\frac{(\lambda-1)}{\lambda}\frac{1}{T}\frac{1}{C^{*\frac{1}{\sigma}}}\right]^{\frac{1}{\omega}}. \qquad \text{(III.29)}$$

Combining the steady state versions of (III.24) and (III.25), with (III.27) yields:

$$Y = C^* \left(\frac{P}{P_H}\right)^\theta \left[(1-a)T^\sigma \left(\frac{P_H}{P}\right)^\sigma + aT^\theta \left(\frac{P_H}{P}\right)^\theta\right], \qquad \text{(III.30)}$$

where $\frac{P}{P_H} = \left[(1-a) + aT^{1-\theta}\right]^{\frac{1}{1-\theta}}$ from (III.9). Thus, (III.29) and (III.30) constitute a two equation system for Y and T, with a unique solution given by $T = 1$ and

$$Y = C^* = \left[\frac{\lambda - 1}{\lambda}\right]^{\frac{1}{\omega + \frac{1}{\sigma}}} > 0.$$

Consequently, $P = P_H$, $Q = 1$, and from (III.27), $C^* = C = Y$, and hence trade is balanced in the steady state.

III.2.6 Log-linearization

Since we are concerned with issues of local equilibrium determinacy I focus on a perfect foresight version of the small open economy model (i.e. the expectation operators are dropped), and the equilibrium conditions are log-linearized around the steady state. Without loss of generality it is assumed that foreign consumption, which is exogenous, remains constant at its steady state value, $C_t^* = C^*$. In what follows all hatted variables denote log-deviations from steady state values. Log-linearizing (III.15) yields the IS equation for the small open economy:

$$\widehat{C}_t = \widehat{C}_{t+1} - \sigma\left[\widehat{R}_t - \widehat{\pi}_{t+1}\right], \tag{III.31}$$

and from (III.9) it follows that the CPI inflation rate can be written as:

$$\widehat{\pi}_t = \widehat{\pi}_t^H + a\left(\widehat{T}_t - \widehat{T}_{t-1}\right), \tag{III.32}$$

where $(1-a)\widehat{T}_t = \widehat{Q}_t = \sigma^{-1}\widehat{C}_t$ from log-linearizing (III.17) and (III.25). Log-linearizing the price-setting equations (III.20) and (III.21) yields the *New Keynesian Phillips Curve* (NKPC):

$$\widehat{\pi}_t^H = \beta\widehat{\pi}_{t+1}^H + \kappa\widehat{mc}_t, \tag{III.33}$$

where $\widehat{\pi}_t^H \equiv \widehat{P}_{H,t} - \widehat{P}_{H,t-1}$ is producer price inflation (PPI) and $\kappa \equiv \frac{(1-\psi)(1-\beta\psi)}{\psi} > 0$ is the real marginal cost elasticity of inflation. Log-linearizing the aggregate production function (III.26) yields: $\widehat{Y}_t = \widehat{L}_t$.[6] Consequently, real marginal cost can be expressed as:

$$\widehat{mc}_t = \omega\widehat{Y}_t + \frac{1}{\sigma}\widehat{C}_t + a\widehat{T}_t = \frac{\Lambda_1}{\sigma(1-a)}\widehat{C}_t, \tag{III.34}$$

[6] As shown by Gali and Monacelli (2005), price dispersion deviations around the steady state are of second order. Thus $\widehat{d}_t = 0$.

where $\Lambda_1 \equiv 1 + \omega\sigma + \omega a(2 - a)(\theta - \sigma) > 0$.[7] The local dynamics of the small open economy is summarized by the dynamic system given by the equilibrium conditions (III.31)–(III.34), along with the log-linearized version of the monetary policy rule (III.22).

III.3 Determinacy analysis for open economies

I now study the determinacy properties of the baseline small open economy model under alternative specifications for the monetary policy rule. In each case I examine the necessary and sufficient conditions under which the monetary policy rule ensures local uniqueness of equilibrium.

III.3.1 How important is trade openness?

Lets start by considering the determinacy implications of a monetary policy rule that reacts to forward-looking CPI inflation. Writing (III.22) in log-linear form yields $\widehat{R}_t = \mu\widehat{\pi}_{t+1}$, and the log-linearized model can be reduced to the following two-dimensional system:

$$\begin{bmatrix} \widehat{C}_{t+1} \\ \widehat{\pi}^H_{t+1} \end{bmatrix} = \begin{bmatrix} 1 - \frac{(\mu-1)\kappa\Lambda_1}{\beta(1-a\mu)} & \frac{(\mu-1)(1-a)\sigma}{\beta(1-a\mu)} \\ -\frac{\kappa\Lambda_1}{\beta\sigma(1-a)} & 1/\beta \end{bmatrix} \begin{bmatrix} \widehat{C}_t \\ \widehat{\pi}^H_t \end{bmatrix}.$$

Proposition 1. *If the monetary policy rule reacts to forward-looking CPI inflation, the necessary and sufficient condition for equilibrium determinacy is:*

$$1 < \mu < 1 + \frac{2(1-a)(1+\beta)}{\kappa\Lambda_1 + 2a(1+\beta)} \equiv \Gamma_1. \qquad \text{(III.35)}$$

Proof. The determinant of the coefficient matrix A is $\det(A) = 1/\beta > 1$. Since there is no predetermined variable, determinacy requires that both eigenvalues are outside the unit circle, which is satisfied if and only if $1 + \det(A) - \text{tr}(A) > 0$, and $1 + \det(A) + \text{tr}(A) > 0$, where tr(A) denotes the trace of the coefficient matrix A. The first inequality holds if and only if $1 < \mu < 1/a$, and the second inequality requires $\mu < \Gamma_1$. Since $\Gamma_1 < 1/a$, the upper bound $1/a$ is redundant. □

Note that for a closed economy (i.e. $a \to 0$) the determinacy condition (III.35) collapses to:

$$1 < \mu < 1 + \frac{2(1+\beta)}{\kappa(1+\omega\sigma)} \equiv \Gamma^c. \qquad \text{(III.36)}$$

[7] By combining the log-linearized versions of equations (III.9) and (III.24), output can be expressed as: $\widehat{Y}_t = (1-a)\widehat{C}_t + a\theta\widehat{Q}_t + a\theta\widehat{T}_t$.

Proposition 1 suggests that while the Taylor principle ($\mu > 1$) is a necessary condition for equilibrium determinacy, it is not sufficient. It is straightforward to verify that the upper bound on the inflation response coefficient is lower for open economies than closed economies $\Gamma_1 < \Gamma^c$, and that $\partial\Gamma_1/\partial\psi > 0$ and $\partial\Gamma_1/\partial a < 0$. This mirrors the key insight of De Fiore and Liu (2005): indeterminacy becomes more severe, as the degree of openness to international trade increases.

Baseline parameterization. The quantitative size of the determinacy region is remarkably small for plausible parameter values. To illustrate this I employ the following baseline parameterization. As justified in Woodford (2003), I set the discount factor $\beta = 0.99$, the labor supply elasticity $\omega = 0.47$, and the intertemporal substitution elasticity of consumption $\sigma = 6.37$. Following Taylor (1999), I set $\psi = 0.75$, which constitutes an average price duration of one year. Finally, the trade price elasticity is set at $\theta = 1$ consistent with macro-level studies (see, e.g., Bergin, 2006).

Figure III.1 illustrates the regions of (in)determinacy for combinations of the inflation response coefficient μ and the degree of trade openness a. In addition, I also plot the (in)determinacy regions for alternative values of $\psi = 0.67, 0.80, 0.83$, which imply that prices are fixed on average for three, five, and six quarters, respectively. By inspection of Figure III.1, the degree of trade openness exerts a destabilizing effect on the perfect foresight equilibrium. The upper bound on μ is tighter, the more open is the economy to international trade, and the lower the degree of price stickiness. For example, with $\psi = 0.75$ the upper bound on μ for a closed economy is $\Gamma^c \approx 12.61$, whereas with a plausible degree of trade openness $a = 0.4$,[8] the upper bound is dramatically reduced $\Gamma_1 \approx 2.33$. While a higher degree of price stickiness increases significantly the upper bound for closed economies, the increase in magnitude is significantly dampened for open economies. For example, with $\psi = 0.83$ and $a = 0$ the upper bound becomes $\Gamma^c \approx 28.29$, whereas with $a = 0.4$ the upper bound is $\Gamma_1 \approx 2.42$.

Why does openness to international trade have such a destabilizing effect on the economy? The important thing to note is that the future CPI inflation rate depends on both the rate of future PPI inflation and future adjustments in the terms of trade:

$$\widehat{\pi}_{t+1} = \widehat{\pi}^H_{t+1} + a\left(\widehat{T}_{t+1} - \widehat{T}_t\right). \qquad \text{(III.37)}$$

Under the Taylor principle, an increase in inflationary expectations $\uparrow \widehat{\pi}_{t+1}$ raises both the nominal and the real interest rate. The latter reduces aggregate demand,

[8] See Table I.2 (p. 33).

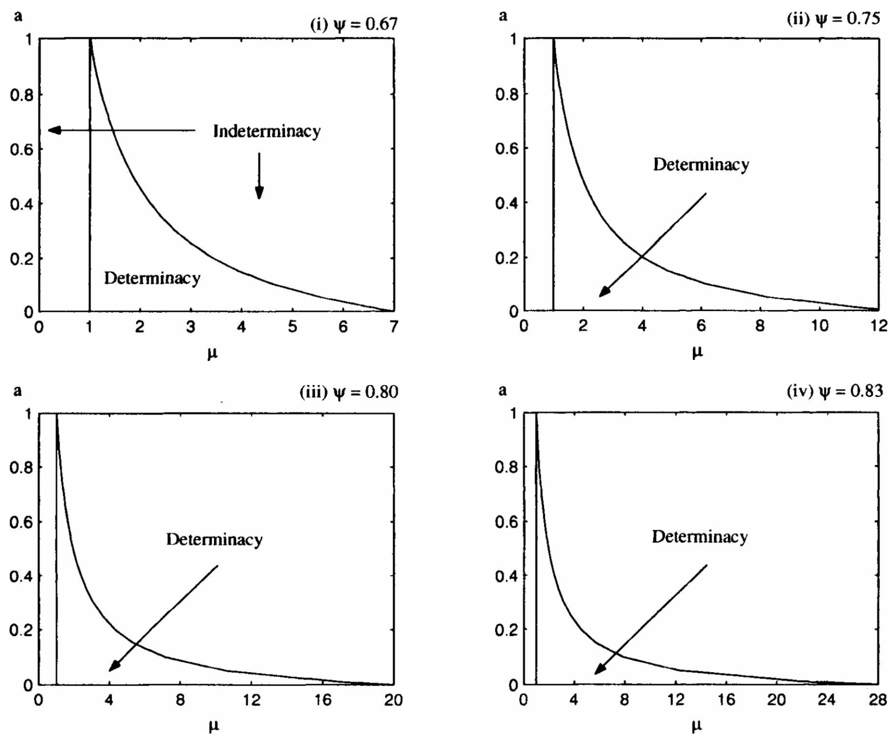


Figure III.1. Regions of (in)determinacy under a forward-looking feedback rule

which decreases future marginal cost, and via the next-period NKPC (III.33), exerts downward pressure on *future PPI inflation* $\downarrow \hat{\pi}^H_{t+1}$. *In a closed economy, since* PPI and CPI inflation are the same concept, the initial inflationary expectation cannot be validated for empirically plausible values of μ. However, in open economies the increase in the real interest rate also results in a current improvement in the terms of trade $\downarrow \hat{T}_t$. From (III.37) this *trade channel of monetary policy* generates upward pressure on the future CPI inflation rate. Consequently, the initial increase in inflationary expectations can be self-fulfilling, despite future PPI inflation falling $\downarrow \hat{\pi}^H_{t+1}$. As the degree of trade openness determines the weight of influence of terms of trade adjustments on CPI inflation, the higher the degree of trade openness, the more prone is the economy to indeterminacy.

It is essential to point out the importance of forward-looking monetary policy in generating the previous finding. Suppose that the interest rate reacts to contemporaneous inflation rather than forward-looking inflation: i.e. $\hat{R}_t = \mu\hat{\pi}_t$. Now the log-linearized model reduces to the following three-dimensional system:

$$\mathbf{x}_{t+1}^1 = \mathbf{A}_1 \mathbf{x}_t^1, \qquad \mathbf{x}_t^1 = \left[\widehat{C}_t \; \widehat{\pi}_t^H \; \widehat{C}_{t-1}\right]',$$

$$\mathbf{A}_1 \equiv \begin{bmatrix} 1 + a\mu + \kappa\Lambda_1\beta^{-1} & \sigma(1-a)\left[\mu - \beta^{-1}\right] & -a\mu \\ -\frac{\kappa\Lambda_1}{\beta\sigma(1-a)} & 1/\beta & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Proposition 2. *If the monetary policy rule reacts to current-looking CPI inflation, the necessary and sufficient conditions for equilibrium determinacy are $\mu > 1$, and either:*

$$\frac{2\beta - 1 - \kappa\Lambda_1}{a\beta} < \mu, \quad or \quad 1 - \beta + \mu(a\beta + \kappa\Lambda_1) > \frac{a\mu}{\beta}\left[1 + \kappa\Lambda_1 - a\mu(1-\beta)\right]. \tag{III.38}$$

Proof. With one predetermined variable $\widehat{C}_{t-1}$, determinacy requires that $\mathbf{A}_1$ has exactly one eigenvalue inside the unit circle, and two eigenvalues outside the unit circle. The three eigenvalues of $\mathbf{A}_1$ are solutions to the cubic equation $r^3 + a_2 r^2 + a_1 r + a_0 = 0$, where $a_2 = -1 - 1/\beta - a\mu - \kappa\Lambda_1\beta^{-1}$, $a_1 = 1/\beta + a\mu(1 + 1/\beta) + \kappa\Lambda_1\mu\beta^{-1}$, and $a_0 = -a\mu\beta^{-1}$. Since $-1 + a_2 - a_1 + a_0 < 0$, the equilibrium is determinate if and only if: (i) $1 + a_2 + a_1 + a_0 > 0$, and (ii) either $|a_2| > 3$, or $a_0^2 - a_0 a_2 + a_1 - 1 > 0$. The first inequality requires $\mu > 1$, and the final two inequalities yield (III.38). □

With a current-looking rule, provided at least one of the conditions given in (III.38) is satisfied, then the Taylor principle ensures a unique equilibrium independent of the degree of trade openness. The numerical analysis suggests that the second condition given in (III.38) is always satisfied for standard parameter values. This becomes transparent by considering the case where $\beta \to 1$, since this condition collapses to $1 > a$.

Linnemann and Schabert (2006) where the first to highlight that trade openness only exerts a destabilizing effect on the economy if the interest rate rule is forward-looking. The intuition behind their result is as follows. Under a contemporaneous policy rule, an increase in inflationary expectations $\uparrow \widehat{\pi}_{t+1}$, increases current inflation $\uparrow \widehat{\pi}_t$, and results in a rise in the real interest rate under the Taylor principle. This reduces aggregate demand, which via the NKPC (III.33), reduces current PPI inflation $\downarrow \widehat{\pi}_t^H$, thereby exerting downward pressure on current CPI inflation via (III.32):

$$\downarrow \widehat{\pi}_t = \downarrow \widehat{\pi}_t^H + a\left(\widehat{T}_t - \widehat{T}_{t-1}\right),$$

where $\widehat{T}_{t-1}$ is predetermined. In open economies, the increase in the real interest rate also results in an improvement in the terms of trade $\downarrow \widehat{T}_t$, which exerts additional downward pressure on CPI inflation. Consequently, self-fulfilling inflationary expectations cannot be supported under a current-looking feedback rule.

III.3.2 PPI inflation vs. CPI inflation

Inspired by the empirical evidence I have only considered monetary policy rules that respond to the CPI inflation rate. However, central banks could alternatively conduct monetary policy by reacting to the PPI inflation rate. Suppose that the monetary authority now adjusts the interest rate according to the policy rule: $\widehat{R}_t = \mu\widehat{\pi}^H_{t+1}$. The log-linearized model reduces to the following two-dimensional system:

$$\begin{bmatrix} \widehat{C}_{t+1} \\ \widehat{\pi}^H_{t+1} \end{bmatrix} = \begin{bmatrix} 1-(\mu-1)\kappa\Lambda_1\beta^{-1} & (\mu-1)(1-a)\sigma\beta^{-1} \\ -\frac{\kappa\Lambda_1}{\beta\sigma(1-a)} & 1/\beta \end{bmatrix} \begin{bmatrix} \widehat{C}_t \\ \widehat{\pi}^H_t \end{bmatrix}.$$

Proposition 3. *If the monetary policy rule reacts to forward-looking PPI inflation, the necessary and sufficient condition for equilibrium determinacy is:*

$$1 < \mu < 1 + \frac{2(1+\beta)}{\kappa\left[1+\omega\sigma+\omega a(2-a)(\theta-\sigma)\right]} \equiv \Gamma_2. \qquad \text{(III.39)}$$

Proof. Since there is no predetermined variable, and the determinant of the coefficient matrix A is $\det(A) = 1/\beta > 1$, determinacy requires $1+\det(A)-\text{tr}(A) > 0$, and $1+\det(A)+\text{tr}(A) > 0$, where $\text{tr}(A)$ denotes the trace of A. The first inequality holds if and only if $\mu > 1$, and the second inequality holds if and only if $\mu < \Gamma_2$. □

Since $\Gamma_1 < \Gamma_2$, Proposition 3 suggests that reacting to PPI inflation significantly reduces the possibility of indeterminacy. Indeed, for the baseline parameterization given in Section III.3.1 open economies are no more susceptible to indeterminacy than closed economies. Analytically, this is easily observed by comparing the upper bounds given in (III.36) and (III.39), where $\Gamma_2 < \Gamma^c$, only if $\theta > \sigma$. However, even in this case determinacy is still easily achievable. For example, if I follow Llosa and Tuesta (2008) and set $\theta = 1.5$ and $\sigma = 0.2$, this generates an upper bound on μ of $\Gamma_2 \approx 32.22$.[9] This compares to an upper bound of $\Gamma_1 \approx 2.38$

[9] The remaining parameters β, ω and ψ are unchanged from the baseline parameterization. I have set $a = 0.4$ for illustrative purposes.

if CPI inflation is the inflation indicator used in the feedback rule. Consequently, indeterminacy is no longer a practical problem for open economies if the monetary policy rule reacts to PPI inflation. This finding coincides with the small open economy studies of Linnemann and Schabert (2006) and Llosa and Tuesta (2008), and is the basis for the literature advocating that policymakers should switch from their current widespread use of CPI inflation, towards PPI inflation, in the setting of monetary policy.

III.3.3 Reacting to the exchange rate

Given the indeterminacy problem identified under a forward-looking CPI inflation feedback rule (III.22), a natural question to ask is whether it is beneficial for monetary authorities to explicitly react to exchange rate fluctuations? Suppose the central bank adjusts the nominal interest rate according to the following policy rule:

$$\widehat{R}_t = \mu_\pi \widehat{\pi}_{t+1} + \mu_e \Delta \widehat{e}_t, \tag{III.40}$$

where $\mu_\pi, \mu_e \geq 0$, and $\Delta \widehat{e}_t \equiv \widehat{e}_t - \widehat{e}_{t-1}$.[10] Using the baseline parameter values of Section III.3.1, Figure III.2 illustrates the regions of (in)determinacy for four alternative values of the nominal exchange rate coefficient: $\mu_e = 0, 0.1, 0.6, 1.0$. By responding to fluctuations in the nominal exchange rate this reduces the lower bound on the inflation response coefficient required for determinacy $\mu_\pi > 1 - \mu_e$, and increases the upper bound Γ_1 given in (III.35) by magnitude $\frac{2\mu_e(1+\beta)+\kappa\Lambda_1\mu_e}{\kappa\Lambda_1+2a(1+\beta)}$. Why does reacting to the nominal exchange rate help to induce determinacy? The inclusion of the nominal exchange rate in the policy rule crucially implies a policy response to the lagged interest rate. To see this, first note that with complete asset markets the model implies the following (log-linearized) uncovered interest parity (UIP) condition: $\widehat{R}_t = \Delta \widehat{e}_{t+1}$. Using the UIP condition I can rewrite the policy rule (III.40) as: $\widehat{R}_t = \mu_\pi \widehat{\pi}_{t+1} + \mu_e \widehat{R}_{t-1}$. Consequently, monetary policy now also responds to current CPI inflation, which as previously discussed in Section III.3.1 exerts a stabilizing effect on the economy. Hence, by reacting to fluctuations in the nominal exchange rate this results in policy inertia or interest-rate smoothing, thereby helping to reduce the indeterminacy problem.

[10] The monetary policy rule can alternatively be designed to respond to fluctuations in the real exchange rate. I postpone an examination of real exchange rate rules until Section III.4 below. However, the conclusions of this section remain unaffected by the replacement of the nominal exchange rate with the real exchange rate in the feedback rule.

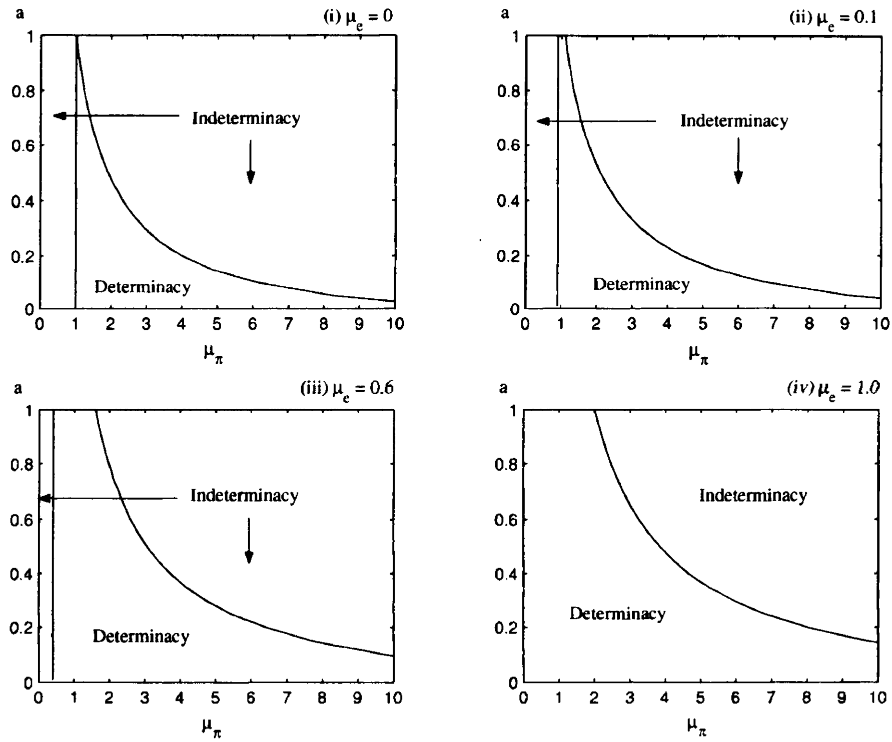


Figure III.2. Regions of (in)determinacy when the feedback rule also reacts to current changes in the nominal exchange rate

Note, however, that how the exchange rate enters the policy rule crucially matters. To see this, now suppose that the monetary policy rule is given by:[11]

$$\widehat{R}_t = \mu_\pi \widehat{\pi}_{t+1} + \mu_e \Delta \widehat{e}_{t+1}.$$

The log-linearized model reduces to the following two-dimensional system:

$$\begin{bmatrix} \widehat{C}_{t+1} \\ \widehat{\pi}^H_{t+1} \end{bmatrix} = \begin{bmatrix} 1 - \frac{(\mu_\pi+\mu_e-1)\kappa\Lambda_1}{\beta(1-a\mu_\pi-\mu_e)} & \frac{(\mu_\pi+\mu_e-1)(1-a)\sigma}{\beta(1-a\mu_\pi-\mu_e)} \\ -\frac{\kappa\Lambda_1}{\beta\sigma(1-a)} & 1/\beta \end{bmatrix} \begin{bmatrix} \widehat{C}_t \\ \widehat{\pi}^H_t \end{bmatrix}.$$

Proposition 4. *If the monetary policy rule reacts to forward-looking CPI inflation and forward-looking changes in the nominal exchange rate, the necessary and sufficient condition for equilibrium determinacy is:*

[11] Lubik and Schorfheide (2007) find evidence that both the Bank of Canada and the Bank of England implement a similar policy rule.

$$1 - \mu_e < \mu_\pi < (1 - \mu_e)\left[1 + \frac{2(1-a)(1+\beta)}{\kappa\Lambda_1 + 2a(1+\beta)}\right] \equiv \Gamma_3.$$

Proof. Since there is no predetermined variable, and the determinant of the coefficient matrix A is $\det(A) = 1/\beta > 1$, determinacy requires $1+\det(A)-\mathrm{tr}(A) > 0$, and $1+\det(A)+\mathrm{tr}(A) > 0$, where tr(A) denotes the trace of A. The first inequality holds if and only if $1 - \mu_e < \mu_\pi < (1 - \mu_e)/a^{-1}$, and the second inequality holds if and only if $\mu_\pi < \Gamma_3$. Note that for $\mu_\pi \geq 0$, $\Gamma_3 < (1 - \mu_e)/a^{-1}$, and thus the upper bound $(1 - \mu_e)/a^{-1}$ is redundant. □

In this case, by reacting to future nominal exchange rate changes the central bank now exacerbates the indeterminacy problem. By inspection of Γ_3, any $\mu_e > 1$ makes determinacy impossible. For values of $\mu_e < 1$, it follows that $\Gamma_3 < \Gamma_1$, and thus reacting to future exchange rate fluctuations reduces the upper bound on the inflation response coefficient. Using the baseline parameterization of Section III.3.1, and setting the degree of trade openness $a = 0.4$, then for any $\mu_e \geq 0.58$ determinacy is not possible under the Taylor principle. Now determinacy is only possible under a passive monetary policy (i.e. $\mu < 1$). For example, if $\mu_e = 0.6$ then determinacy is obtained within the interval $0.4 < \mu_\pi < 0.932$.

III.4 THE IMPORTANCE OF MONEY AND CAPITAL

So far the analysis has followed the existing small open economy literature and ignored both the role of money and investment spending from the modeling framework. However, as shown by McKnight (2011a, 2011b) for two-country frameworks the assumptions of a cashless economy and a labor-only economy are not harmless. In Section III.4.1 below, money is introduced into the baseline model, and then in Section III.4.2, I allow for capital and investment spending. As will become clear, not only are the determinacy properties significantly altered in the presence of either money or capital, but the key policy recommendations of the previous section are reversed.

III.4.1 Money matters

Money is introduced into the baseline model by assuming that domestic real money balances enter into the utility function of the representative household:[12,13]

[12] I adopt a money-in-the-utility function approach because of its popularity. However, employing a cash-in-advance approach does not alter the conclusions presented here.

[13] I focus on a perfect foresight version of the model.

$$\max \sum_{t=0}^{\infty} \beta^t \left[U(C_t) + V(A_t/P_t) - H(L_t)\right]$$

where:

$$A_t \equiv M_t + X_t^m + B_t - \frac{B_{t+1}}{R_t}.$$

In the above, $X_t^m \equiv M_t^s - M_{t-1}^s$ denotes the growth in the money supply, which is transferred to the household from the monetary authority. Here, it is assumed that the stock of money that yields utility is the stock of money holdings after bonds have been purchased in the financial markets but before income has been received or consumption goods have been purchased. Carlstom and Fuerst (2001) call this *cash-in-advance timing*, since what generates transactions services is the money balances held in advance of goods trading. A common (but less realistic) alternative is that end-of-period money balances enter into the utility function (i.e. $A_t = M_{t+1}$). Here, what facilitates transaction services is the amount of money remaining *after* goods market trading. In this case, provided the utility function is separable between consumption and real money balances, the determinacy conditions of the monetary economy mirror the cashless economy.[14,15]

With money, the modified consumption Euler equation is given by:

$$\beta \frac{U_c\left(C_{t+1}\right)}{P_{t+1}} = \frac{U_c\left(C_t\right)}{P_t} \frac{1}{R_{t+1}}, \tag{III.41}$$

which is the same as the cashless version (III.15) except that the nominal interest rate is scrolled forward one-period. With money, there is an additional first-order condition for optimal money holdings:

$$\frac{V_m(m_t)}{U_c(C_t)} = R_t - 1, \tag{III.42}$$

and equation (III.17) is subsequently amended:

$$Q_t = q_0^m \frac{U_{c^*}(C_t^*) + V_{m^*}(m_t^*)}{U_c(C_t) + V_m(m_t)} = q_0^m \frac{U_{c^*}(C_t^*)R_t^*}{U_c(C_t)R_t}, \tag{III.43}$$

[14] See McKnight (2011b) for further details.

[15] Employing a two-country framework, McKnight and Mihailov (2012) investigate the determinacy implications of alternative monetary policy rules assuming end-of-period money balances and non-separability of the utility function. They show that in the presence of empirically plausible real balance effects indeterminacy is a serious problem when the policy rule reacts to CPI inflation, which cannot be alleviated by reacting to the exchange rate.

where I have used the money demand equation (III.42), and its foreign equivalent, to eliminate $V_m(m_t)$ and $V_{m^*}(m_t^*)$, respectively. The constant q_0^m is defined as: $q_0^m \equiv Q_0 \left[\frac{U_c(C_0)+V_m(m_0)}{U_{c^*}(C_0^*)+V_{m^*}(m_0^*)}\right]$. Comparing (III.17) and (III.43) highlights the second difference between the monetary and cashless economies. In a cashless economy the real exchange rate is equated with the marginal utilities of consumption, whereas in a monetary economy the marginal utilities of money are also included. This reflects the fact that a bond sale for consumption purposes increases the utility from current consumption and current liquidity.

The log-linearized equilibrium system is now given by the modified IS equation:

$$\widehat{C}_t = \widehat{C}_{t+1} - \sigma \left[\widehat{R}_{t+1} - \widehat{\pi}_{t+1}\right], \tag{III.44}$$

and real marginal cost can be expressed as:

$$\widehat{mc}_t = \omega \widehat{Y}_t + \frac{1}{\sigma}\widehat{C}_t + a\widehat{T}_t = \frac{\Lambda_1}{\sigma(1-a)}\widehat{C}_t - \left(\frac{a}{1-a} + \frac{a\omega\theta(2-a)}{1-a}\right)\widehat{R}_t, \tag{III.45}$$

since now $(1-a)\widehat{T}_t = \widehat{Q}_t = \sigma^{-1}\widehat{C}_t - \widehat{R}_t$. Equations (III.32) and (III.33) remain unchanged.

To see the importance of money lets first suppose that the central bank adopts the following monetary policy rule: $\widehat{R}_t = \mu\widehat{\pi}_t^H$. This case was previously analyzed by McKnight (2011*b*) in the context of a two-county model. The log-linearized model with money reduces to the following two-dimensional system:

$$\begin{bmatrix} \widehat{C}_{t+1} \\ \widehat{\pi}_{t+1}^H \end{bmatrix} = \begin{bmatrix} 1 - \frac{[\mu-(1-a)]\kappa\Lambda_1}{\beta(1-a)} & [\mu-(1-a)]\sigma\beta^{-1}\Lambda_m - a\sigma\mu \\ -\frac{\kappa\Lambda_1}{\beta\sigma(1-a)} & \Lambda_m\beta^{-1} \end{bmatrix} \begin{bmatrix} \widehat{C}_t \\ \widehat{\pi}_t^H \end{bmatrix},$$

where $\Lambda_m \equiv 1 + a\kappa\mu(1-a)^{-1}\left[1 + \omega\theta(2-a)\right]$.

Proposition 5. *If the monetary policy rule reacts to current-looking PPI inflation, the necessary and sufficient conditions for equilibrium determinacy in a monetary economy are:*

Case I: $\mu > 1$ *if* $a\omega\theta(2-a) - \omega\sigma(1+a)(1-a) > 1$, *or*

$$1 < \mu < \min\left\{\Gamma_4^A, \Gamma_4^B\right\} \quad \textit{if} \quad a\omega\theta(2-a) - \omega\sigma(1+a)(1-a) < 1, \tag{III.46}$$

Case II: $\Gamma_4^A < \mu < 1$, *and* $a\omega\theta(2-a) - \omega\sigma(1+a)(1-a) < 1$,

where $\Gamma_4^A \equiv \frac{2(1+\beta)+\kappa\Lambda_1}{\kappa[1+\omega\sigma(1+a)(1-a)-a\omega\theta(2-a)]}$; $\Gamma_4^B \equiv \frac{(1-\beta)}{a\kappa\omega[(2-a)(\sigma-\theta)-\sigma]}$.

Proof. The two eigenvalues are solutions to the quadratic equation $r^2+a_1 r+a_0 = 0$, where $a_1 = \text{tr(A)}$, and $a_0 = \det(\text{A})$. Since there is no predetermined variable both eigenvalues must be outside the unit circle for determinacy, which requires that one of the following two cases is satisfied. Case I: $a_0 > 1$, $1 + a_0 + a_1 > 0$, $1 + a_0 - a_1 > 0$; Case II: $1 + a_0 + a_1 < 0$, $1 + a_0 - a_1 < 0$. First note that $\mu > 1$ always satisfies Case I if $a\omega\theta(2-a) - \omega\sigma(1+a)(1-a) > 1$. Otherwise, the three inequalities can be reduced to (III.46). Case II requires $a\omega\theta(2-a)-\omega\sigma(1+a)(1-a) < 1$, and $\Gamma_4^A < \mu < 1$. □

For a closed economy, the determinacy condition is given by (III.36) with upper bound Γ^c. Consequently, Proposition 5 suggests that whether indeterminacy is greater or smaller for open economies depends on the relative size of a, ω, θ, and σ. In order to gain some further insight, a numerical analysis is carried out using the baseline parameter values of Section III.3.1. Note that under this parameterization $a\omega\theta(2-a) - \omega\sigma(1+a)(1-a) < 1$, for all permissable values of a, and Case II never applies. Thus, (III.46) gives the empirically relevant determinacy condition. The numerical analysis suggests that $\Gamma_4^B < \Gamma_4^A$, which significantly constrains the ability of the Taylor principle to achieve determinacy. For example, with $\psi = 0.75$ the upper bound on μ for a closed economy is $\Gamma^c \approx 12.61$, whereas with only a small degree of trade openness $a = 0.05$ the upper bound is dramatically reduced $\Gamma_4^B \approx 1.21$. Indeed, determinacy is impossible for any μ within the interval $0.06 < a < 0.76$.[16] This is in stark contrast to a rule that reacts to CPI inflation, $\hat{R}_t = \mu\hat{\pi}_t$, where the range of determinacy is significantly larger.[17] For example, with $a = 0.05$ the upper bound increases dramatically $\mu < 40.50$. For higher degrees of trade openness, the upper bound becomes so large that the Taylor principle effectively holds. Therefore, in stark contrast to Section III.3.2, in a monetary economy CPI inflation can be preferable to PPI inflation in the setting of monetary policy.

The intuition behind these findings are as follows. In a monetary economy marginal cost (III.45), and consequently the dynamics of PPI inflation (III.33), now depend on the nominal interest rate, which enters as a *negative cost shock*.

[16] The upper bound Γ_4^B ceases to apply when $(2-a)(\sigma-\theta) \leq \sigma$, which for the baseline parameterization requires a very high degree of trade openness $a \geq 0.81$.

[17] See McKnight (2011b) for an analytical derivation of the determinacy conditions under this policy rule.

Therefore, if $\widehat{R}_t = \mu\widehat{\pi}_t^H$ this makes self-fulfilling expectations more likely. However, from (III.32) the behavior of CPI inflation depends on both the PPI inflation rate and the terms of trade. Since increases in the real interest rate result in an improvement in the terms of trade $\downarrow \widehat{T}_t$, this generates downward pressure on the CPI inflation. It follows that monetary policy, by reacting to CPI inflation, can help to prevent self-fulfilling inflationary expectations by offsetting the negative cost shock to PPI inflation.

III.4.2 Introducing investment

So far capital and investment spending have been ignored from the baseline model. However, since investment spending is a very important component of aggregate demand, its inclusion will have significant implications for the transmission mechanism of monetary policy. In order to introduce capital into the baseline model it is assumed that there is an economy-wide rental market for the capital stock.[18] The changes to the baseline model are briefly outlined below.

Firms. To produce output intermediate firms hire labor and rent capital from the representative household, given the real wage rate w_t and the capital rental cost rr_t. Suppose a firm of type i has the following production technology:

$$y_t(i) = K_t(i)^\alpha L_t(i)^{1-\alpha}, \tag{III.47}$$

where K and L represent capital and labor usage respectively, and the input share is $0 < \alpha < 1$. Given competitive prices of labor and capital, cost-minimization yields:

$$w_t = mc_t(1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)\left(\frac{K_t(i)}{L_t(i)}\right)^\alpha, \tag{III.48}$$

$$rr_t = mc_t\alpha\left(\frac{P_{H,t}}{P_t}\right)\left(\frac{L_t(i)}{K_t(i)}\right)^{1-\alpha}. \tag{III.49}$$

The firms price-setting problem remains unchanged.

Households. The representative household owns the capital stock K and makes all investment decisions I according to the following law of motion:

$$K_{t+1} = (1-\delta)K_t + I_t, \tag{III.50}$$

[18] An alternative approach is to assume that capital is firm-specific. For further details see Sveen and Weinke (2005) and Duffy and Xiao (2011).

where $0 < \delta < 1$ is the depreciation rate of capital. Consequently, there is an additional first-order condition for optimal household investment:

$$U_c(C_t) = \beta U_c(C_{t+1}) \left[rr_{t+1} + (1-\delta) \right]. \tag{III.51}$$

Finally, the market clearing condition (III.24) is modified:

$$Y_t = \left(\frac{P_{H,t}}{P_t} \right)^{-\theta} \left[(1-a)Z_t + aZ_t^* Q_t^{\theta} \right] \tag{III.52}$$

where $Z_t = C_t + I_t$.

In what follows it will be useful to define the labor-capital ratio $V_t \equiv \frac{L_t}{K_t}$. Furthermore, for analytical simplicity it is assumed that the labor supply elasticity is infinite (i.e. $\omega = 0$). Equations (III.31)–(III.33) of the log-linearized model are unchanged in the presence of investment activity. However, the marginal cost condition (III.34) is now expressed as:

$$\widehat{mc}_t = \frac{1}{\sigma}\widehat{C}_t + a\widehat{T}_t + \alpha\widehat{V}_t = \frac{\widehat{C}_t}{\sigma(1-a)} + \alpha\widehat{V}_t, \tag{III.53}$$

and there are two additional log-linearized equations for the two additional endogenous variables $\widehat{V}_t$ and $\widehat{K}_t$. Using the log-linearized versions of (III.49) and (III.51) generates:

$$\widehat{C}_{t+1} = \widehat{C}_t + \sigma\Lambda_2\widehat{rr}_{t+1} = \widehat{C}_t + \sigma\Lambda_2\left[\widehat{mc}_{t+1} - \frac{a}{\sigma(1-a)}\widehat{C}_{t+1} + (1-\alpha)\widehat{V}_{t+1}\right], \tag{III.54}$$

where $\Lambda_2 \equiv 1 - \beta(1-\delta) > 0$, and combining the aggregate log-linearized version of (III.47), with (III.50) and (III.52) yields:

$$\widehat{K}_{t+1} = \left[1 - \delta + \frac{\delta}{(1-a)s_I}\right]\widehat{K}_t + \frac{\delta}{s_I}\frac{(1-\alpha)}{(1-a)}\widehat{V}_t - \frac{\delta}{s_I}\left[\frac{\theta a(2-a)}{\sigma(1-a)^2} + s_c\right]\widehat{C}_t, \tag{III.55}$$

where $s_I, s_c \in (0, 1)$ are the respective steady state output shares of investment and consumption.

To see the importance of capital and investment spending lets first suppose that the monetary authority adjusts the interest rate according to the policy rule: $\widehat{R}_t = \mu\widehat{\pi}_t^H$. This case was previously analyzed by McKnight (2011*a*) in the context of a two-county model.[19] The log-linearized model with capital and investment spending can be reduced to the following four-dimensional system:

[19] I only consider a contemporaneous specification for the monetary policy rule, since it is now well established from the closed economy literature that determinacy is nearly impossible with forward-looking Taylor-type rules. See, e.g., Carlstrom and Fuerst (2005), Kurozumi and Van Zandweghe (2008), Huang et al. (2009), Duffy and Xiao (2011).

$$\mathbf{x}_{t+1} = \mathbf{A}_2\mathbf{x}_t, \quad \mathbf{x}_t = \left[\widehat{mc}_t \; \widehat{V}_t \; \widehat{\pi}_t^H \; \widehat{K}_t\right]',$$

$$\mathbf{A}_2 \equiv \begin{bmatrix} 1-\alpha(1-a)+\frac{\kappa}{\beta}J_1 & \alpha^2(1-a)-\alpha & \left(\mu-\frac{1}{\beta}\right)J_1 & 0 \\ -(1-a)+\frac{\kappa}{\beta}J_2 & \alpha(1-a) & \left(\mu-\frac{1}{\beta}\right)J_2 & 0 \\ -\frac{\kappa}{\beta} & 0 & \frac{1}{\beta} & 0 \\ -\frac{\delta}{s_I}\left[s_c\sigma(1-a)+\frac{\theta a(2-a)}{1-a}\right] & J_3 & 0 & 1+\delta\left[\frac{1-s_I(1-a)}{s_I(1-a)}\right] \end{bmatrix},$$

where, $J_1 = \left[1+\frac{\alpha(1-\Lambda_2)(1-a)}{\Lambda_2}\right]$; $\quad J_2 = \frac{(1-\Lambda_2)(1-a)}{\Lambda_2}$;

and $\quad J_3 = \frac{\delta}{s_I}\left[\frac{1-\alpha}{1-a}+\frac{\alpha\theta a(2-a)}{1-a}+s_c\sigma\alpha(1-a)\right]$.

Proposition 6. *If the monetary policy rule reacts to current-looking PPI inflation, the necessary and sufficient conditions for equilibrium determinacy with capital and investment spending are $\mu > 1$, and either:*

$$(2\beta-1)\Lambda_2\kappa^{-1} < 1-\beta(1-\delta)[1-(1-a)\alpha], \qquad \text{(III.56)}$$

$$\text{or} \quad \frac{\mu\Lambda_3}{\Lambda_2\beta}\left[\frac{\Lambda_3(\mu-1)}{\Lambda_2}+\Lambda_3-(1+\kappa+\Lambda_2\beta)\right]+(1-\beta)+\kappa\mu+\frac{\Lambda_3}{\Lambda_2} > 0; \qquad \text{(III.57)}$$

where $\Lambda_2 = 1-\beta(1-\delta)$ and $\Lambda_3 = \kappa\alpha(1-a)$.

Proof. By inspection of $\mathbf{A}_2$ one eigenvalue is given by $1+\delta\left[\frac{1-s_I(1-a)}{s_I(1-a)}\right] > 1$, which is outside the unit circle. With one predetermined variable $\widehat{K}_t$, determinacy requires that two of the three remaining eigenvalues lie outside the unit circle, and one eigenvalue is inside the unit circle. The remaining eigenvalues of $\mathbf{A}_2$ are solutions to the cubic equation $r^3+a_2r^2+a_1r+a_0 = 0$, where $a_2 = -1-1/\beta-\kappa J_1\beta^{-1}$; $a_1 = 1/\beta+\kappa J_1\mu\beta^{-1}+\alpha\kappa(1-a)\beta^{-1}\Lambda_2^{-1}$; and $a_0 = -\mu\kappa\alpha(1-a)\beta^{-1}\Lambda_2^{-1}$. Since $-1+a_2-a_1+a_0 < 0$, the equilibrium is determinate if and only if: (i) $1+a_2+a_1+a_0 > 0$, and (ii) either $|a_2| > 3$, or $a_0^2-a_0a_2+a_1-1 > 0$. The first inequality requires $\mu > 1$, and the final two inequalities yields (III.56) and (III.57) respectively. □

The determinacy conditions in Proposition 6 are illustrated using the following parameter values.[20] As in Section III.3.1, I set $\beta = 0.99$. As justified by Sveen

[20] Note that I do not need to assign values for σ or θ since the determinacy conditions are independent of these parameters.

and Weinke (2005), I set the cost share of capital $\alpha = 0.36$ and the depreciation rate of capital $\delta = 0.025$. Following Kurozumi and Van Zandweghe (2008), I set the steady state output share of consumption $s_c = 0.7$ and investment $s_I = 0.3$. For illustrative purposes, I set the degree of trade openness $a = 0.4$. The top, left-hand panel of Figure III.3 illustrates the regions of (in)determinacy for combinations of the inflation response coefficient μ and the degree of price rigidity ψ. By inspection, under the Taylor principle the region of indeterminacy increases, as prices become more sticky, and this indeterminacy region is always greater for the open economy relative to the closed economy.[21] In stark contrast, the top-right panel of Figure III.3 illustrates the regions of (in)determinacy when the policy rule reacts to CPI inflation: $\widehat{R}_t = \mu\widehat{\pi}_t$.[22] The numerical analysis suggests that the Taylor principle easily generates determinacy in this case.

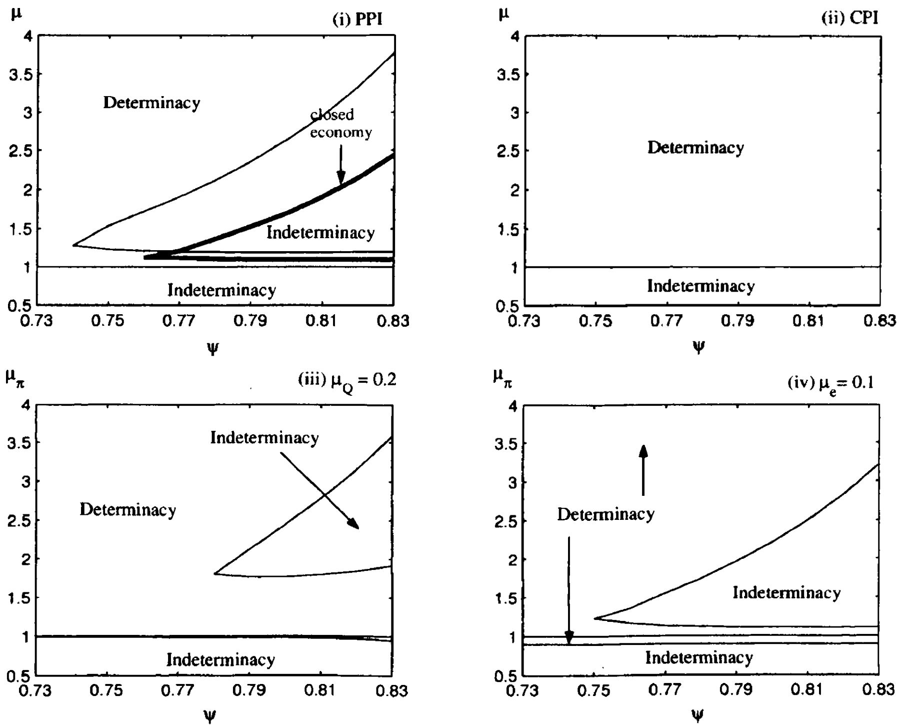


Figure III.3. Regions of (in)determinacy with investment

[21] It can also be shown that the indeterminacy region increases in magnitude as the degree of trade openness increases. See McKnight (2011a) for further details.

[22] See McKnight (2011a) for an analytical derivation of the determinacy conditions under this policy rule.

With capital and investment spending, the above analysis suggests that CPI inflation is preferable to PPI inflation in the setting of monetary policy. This is in stark contrast to the policy recommendations derived in Section III.3.2 using the labor-only model. What is the intuition behind these results? First note that with capital and investment spending indeterminacy arises from the *cost channel of monetary policy*.[23] Since an increase in inflationary expectations results in an increase in the real interest rate under the Taylor principle, from (III.54) this puts upward pressure on the future rental price of capital $\uparrow \widehat{rr}_{t+1}$. This leads to an increase in future marginal cost, higher future PPI inflation $\uparrow \widehat{\pi}^H_{t+1}$, which via the NKPC, implies higher PPI inflation today $\uparrow \widehat{\pi}^H_t$. Consequently, self-fulfilling inflationary expectations can be supported when the policy rule reacts to PPI inflation. However, from (III.32) there are two opposing effects on CPI inflation. While the cost channel generates upward pressure in PPI inflation $\uparrow \widehat{\pi}^H_t$, the trade channel of monetary policy exerts downward pressure on the terms of trade $\downarrow \widehat{T}_t$. Since any increase in PPI inflation can be offset by the trade channel of monetary policy, this reduces the likelihood of indeterminacy if the policy rule reacts to CPI inflation.

The problem of indeterminacy that arises under PPI inflation feedback rules can be mitigated if the policy rule additionally responds to the exchange rate. For example, consider the following monetary policy rule advocated by Taylor (2001): $\widehat{R}_t = \mu_\pi \widehat{\pi}^H_t + \mu_Q \widehat{Q}_t$, where $\mu_\pi, \mu_Q \geq 0$. This leads to a system of the form:

$$\mathbf{x}_{t+1} = \mathbf{A}_3 \mathbf{x}_t, \quad \mathbf{x}_t = \left[\widehat{mc}_t \; \widehat{V}_t \; \widehat{\pi}^H_t \; \widehat{K}_t\right]',$$

$$\mathbf{A}_3 \equiv \begin{bmatrix} 1 - \alpha(1-a) + \left[\frac{\kappa}{\beta} + J_4\right] J_1 & -\alpha[1 - \alpha(1-a)] - \alpha J_1 J_4 & \left(\mu_\pi - \frac{1}{\beta}\right) J_1 & 0 \\ -(1-a) + \left[\frac{\kappa}{\beta} + J_4\right] J_2 & \alpha(1-a)[1 - \mu_Q J_2] & \left(\mu_\pi - \frac{1}{\beta}\right) J_2 & 0 \\ -\frac{\kappa}{\beta} & 0 & \frac{1}{\beta} & 0 \\ -\frac{\delta}{s_I}\left[s_c \sigma(1-a) + \frac{\theta a(2-a)}{1-a}\right] & J_3 & 0 & J_5 \end{bmatrix},$$

where $J_1 = \left[1 + \frac{\alpha(1-\Lambda_2)(1-a)}{\Lambda_2}\right]$; $\quad J_2 = \frac{(1-\Lambda_2)(1-a)}{\Lambda_2}$;

$J_3 = \frac{\delta}{s_I}\left[\frac{1-\alpha}{1-a} + \frac{\alpha\theta a(2-a)}{1-a} + s_c \sigma \alpha(1-a)\right]$; $\quad J_4 = (1-a)\mu_Q$;

and $\quad J_5 = 1 + \delta\left[\frac{1-s_I(1-a)}{s_I(1-a)}\right]$.

[23] Kurozumi and Van Zandweghe (2008) were the first to highlight this channel using a closed economy analysis.

Proposition 7. *If the monetary policy rule reacts to current-looking PPI inflation and the real exchange rate, the necessary and sufficient conditions for equilibrium determinacy with capital and investment spending are:*

$$\kappa(\mu_\pi - 1)[1 - \alpha(1 - a)] + \mu_Q(1 - a)(1 - \beta) > 0, \quad \text{(III.58)}$$

and either:

$$(2\beta - 1)\Lambda_2\kappa^{-1} < 1 - \beta(1 - \delta)[1 - (1 - a)\alpha] + (1 - a)\beta\mu_Q\Lambda_2\kappa^{-1}, \quad \text{(III.59)}$$

or

$$\frac{\mu_\pi\Lambda_3}{\Lambda_2\beta}\left[\frac{\Lambda_3(\mu_\pi - 1)}{\Lambda_2} + \Lambda_3 - (1 + \kappa + \Lambda_2\beta) - \beta(1 - a)\mu_Q\right] + (1 - \beta) + \kappa\mu_\pi + \frac{\Lambda_3}{\Lambda_2} + (1 - a)\mu_Q > 0; \quad \text{(III.60)}$$

where $\Lambda_2 = 1 - \beta(1 - \delta)$ *and* $\Lambda_3 = \kappa\alpha(1 - a)$.

Proof. One eigenvalue is given by $1+\delta\left[\frac{1-s_I(1-a)}{s_I(1-a)}\right] > 1$. With one predetermined variable $\widehat{K}_t$, determinacy requires that two of the remaining three eigenvalues lie outside the unit circle, and one eigenvalue is inside the unit circle. The remaining eigenvalues of $\mathbf{A}_3$ are solutions to the cubic equation $r^3+a_2r^2+a_1r+a_0 = 0$, where $a_2 = -1-1/\beta-\kappa J_1\beta^{-1}-(1-a)\mu_Q$; $a_1 = 1/\beta+\kappa J_1\mu_\pi\beta^{-1}+\alpha\kappa(1-a)\beta^{-1}\Lambda_2^{-1}+(1-a)\mu_Q\beta^{-1}$; and $a_0 = -\mu_\pi\kappa\alpha(1-a)\beta^{-1}\Lambda_2^{-1}$. Since $-1 + a_2 - a_1 + a_0 < 0$, the equilibrium is determinate if and only if: (i) $1 + a_2 + a_1 + a_0 > 0$, and (ii) either $|a_2| > 3$, or $a_0^2 - a_0a_2 + a_1 - 1 > 0$. The first inequality requires (III.58), and the final two inequalities yields (III.59) and (III.60) respectively. □

Proposition 7 suggests that the indeterminacy problem associated with PPI inflation can be mitigated if the policy rule also responds to the real exchange rate. The bottom left-hand panel of Figure III.3 illustrates the (in)determinacy regions when $\mu_Q = 0.2$. By inspection, the area of indeterminacy under the Taylor principle decreases for each value of the degree of price stickiness, ψ. Indeed, the numerical analysis suggests that if $\mu_Q = 0.38$ then indeterminacy is not possible under the Taylor principle for any $\psi \leq 0.83$. The bottom right-hand panel of Figure III.3 also illustrates the implications for (in)determinacy by reacting to current fluctuations in the nominal exchange rate; i.e. $\widehat{R}_t = \mu_\pi\widehat{\pi}_t^H + \mu_e\Delta\widehat{e}_t$, where for illustrative purpose I have set $\mu_e = 0.1$. Overall, the analysis suggests that either nominal or real exchange rate targeting helps to ameliorate the indeterminacy problem when PPI inflation is the inflation indicator used in the policy rule.

The intuition behind these results is as follows. Similar to Section III.3.3, by reacting to the nominal exchange rate this introduces policy inertia into the dynamic system. Now since the policy rule in effect responds to both current and past PPI inflation this helps to prevent indeterminacy. By reacting to the real exchange rate the mechanism for ameliorating the indeterminacy problem is different. Recall that indeterminacy is generated with investment through the cost channel of monetary policy. By inspection of the risk sharing condition $\widehat{Q}_t = \sigma^{-1}\widehat{C}_t$, reacting to the real exchange rate exerts downward pressure on marginal cost through reducing aggregate demand. This weakens the upward pressure on marginal cost generated by the cost channel, thereby making self-fulfilling inflationary expectations less likely.

III.5 Conclusions

The determinacy implications of monetary policy rules are typically examined through the lens of a closed economy. However, a recent literature has suggested that monetary policies appropriate for closed economies may not be sufficient for open economies. Furthermore, open economies face additional issues when designing monetary policy rules. This chapter has attempted to summarize the key findings of the determinacy literature for small open economies. Recent research suggests that open economies can be more susceptible to policy-induced indeterminacy than closed economies. Therefore, policymakers would be ill-advised to blindly follow well-established policy recommendations that ignore the trade channel of monetary policy. However, the policy advice to emerge from this literature is optimistic. For example, while the Taylor principle may not be sufficient to prevent indeterminacy in open economies, policy rules that obey the Taylor principle and give some weight to either the nominal or real exchange rate can help to avoid indeterminacy.

References

Batini, N., Levine, P., Pearlman, J. (2004). Indeterminacy with Inflation-Forecast-Based Rules in a Two-Bloc Model. International Finance Discussion Paper No. 797, Board of Governors of the Federal Reserve System, Washington, DC.

Benhabib, J., Schmitt-Grohé, S., Uribe, M. (2001). Monetary Policy and Multiple Equilibria. *American Economic Review* 91, 167–186.

Bernanke, B., Woodford, M. (1997). Inflation Forecasts and Monetary Policy, Part 2. *Journal of Money, Credit and Banking* 29, 653–684.

Bergin, P. R. (2006). How Well Can the New Open Economy Macroeconomics Explain the Exchange Rate and Current Account? *Journal of International Money and Finance* 25, 675–701.

Bullard, J., Schaling, E. (2009). Monetary Policy, Determinacy and Learnability in a Two-Block World Economy. *Journal of Money, Credit and Banking* 41, 1585–1612.

Calvo, G. (1983). Staggered Prices in a Utility Maximizing Framework. *Journal of Monetary Economics* 12, 383–398.

Carlstrom, C. T., Fuerst, T. S. (2001). Timing and Real Indeterminacy in Monetary Models. *Journal of Monetary Economics* 47, 285–298.

Carlstrom, C. T., Fuerst, T. S. (2005). Investment and Interest Rate Policy: A Discrete Time Analysis. *Journal of Economic Theory* 123, 4–20.

Clarida, R., Galí, J., Gertler, M. (1998). Monetary Policy Rules in Practice: Some International Evidence. *European Economic Review* 42, 1033–1067.

Clarida, R., Galí, J., Gertler, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics* 115, 147–180.

De Fiore, F., Lui, Z. (2005). Does Trade Openness Matter for Aggregate Instability? *Journal of Economic Dynamics and Control* 29, 1165–1192.

Duffy, J., Xiao, W. (2011). Investment and Monetary Policy: Learning and Determinacy of Equilibrium. *Journal of Money, Credit and Banking* 43, 959–992.

Faia, E., Monacelli, T. (2008). Optimal Monetary Policy in a Small Open Economy with Home Bias. *Journal of Money, Credit and Banking* 40, 721–750.

Gali, J., Monacelli, T. (2005). Monetary Policy and Exchange Rate Volatility in a Small Open Economy. *Review of Economic Studies* 72, 707–734.

Huang, K., Meng, Q., Xue, J. (2009). Is Forward-looking Inflation Targeting Destabilizing? The Role of Policy's Response to Current Output under Endogenous Investment. *Journal of Economic Dynamics and Control* 33, 409–430.

Kurozumi, T., Van Zandweghe, W. (2008). Investment, Interest Rate Policy, and Equilibrium Stability. *Journal of Economic Dynamics and Control* 32, 1489–1516.

Leith, C., Wren-Lewis, S. (2009). Taylor Rules in the Open Economy. *European Economic Review* 53, 971–995.

Linnemann, L., Schabert, A. (2006). Monetary Policy and the Taylor Principle in Open Economies. *International Finance* 9, 343–367.

Llosa, G., Tuesta, V. (2008). Determinacy and Learnability of Monetary Policy Rules in Small Open Economies. *Journal of Money, Credit and Banking* 40, 1033–1063.

Lubik, T. A., Marzo, M. (2007). An Inventory of Simple Monetary Policy Rules in a New Keynesian Macroeconomic Model. *International Review of Economics and Finance* 16, 15–36.

Lubik, T. A., Schorfheide, F. (2007). Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation. *Journal of Monetary Economics* 54, 1069–1087.

McKnight, S. (2011a). Investment and Interest Rate Policy in the Open Economy. *Oxford Economic Papers* 63, 673–699.

McKnight, S. (2011b). Should Central Banks Target Consumer or Producer Prices? *International Finance* 14, 445–479.

McKnight, S., Mihailov, A. (2012). Do Real Balance Effects Invalidate the Taylor Principle in Closed and Open Economies? Working Paper No. 10/2012, El Colegio de México.

Mihailov, A. (2006). Operational Independence, Inflation Targeting, and UK Monetary Policy. *Journal of Post Keynesian Economics* 28, 395–421.

Orphanides, A. (2001). Monetary Policy Rules Based on Real-Time Data. *American Economic Review* 91, 964–985.

Orphanides, A. (2004). Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches. *Journal of Money, Credit and Banking* 36, 151–175.

Sveen, T., Weinke, L. (2005). New Perspectives on Capital, Sticky Prices, and the Taylor Principle. *Journal of Economic Theory* 123, 21–39.

Taylor, J. B. (1993). Discretion Versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.

Taylor, J. B. (1999). Staggered Price and Wage Setting in Macroeconomics. In Taylor, J. B., Woodford, M., eds., *Handbook of Macroeconomics*, 1009–1050. Amsterdam: Elsevier North-Holland.

Taylor, J. B. (2001). The Role of the Exchange Rate in Monetary Policy Rules. *American Economic Review* 91, 263–267.

Woodford, M. (2003). *Interest Rates and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ: Princeton University Press.

IV. THE INFORMAL SECTOR IN CONTEMPORARY MODELS OF THE AGGREGATE ECONOMY

JULIO CÉSAR LEAL ORDÓÑEZ[1]
Banco de México

IV.1 INTRODUCTION

In this chapter, I want to review a contemporary branch of the informal sector literature that studies the interaction between the informal sector and aggregate variables. From my perspective, this branch is the most promising area of research on informality, because as opposed to previous work, it analyzes the issue from a wider perspective.

The selection of papers surveyed in this chapter all analyze the issue of informality building answers from the bottom up. That is, a common strategy throughout is to ask how individual decisions are affected by the presence of informality and, in turn, how these distorted individual decisions add-up and impact on aggregate variables. In order to achieve this goal, the group of papers I discuss below all make use of Dynamic General Equilibrium (DGE) models. The goals of this chapter are threefold: 1) to identify the key ingredients in these DGE models; 2) to present the most important differences in the modeling strategies employed in this literature; and 3) to outline the main challenges that this group of authors face.

The traditional literature used to have what La Porta and Shleifer (2008) called "a romantic view" of informality in which the informal sector was that part of the economy where economic activity was being undertaken in the most "pure and efficient" way. This view is now challenged by the modern literature via the use of DGE models. Both views start from the same principle: informal firms do not pay taxes or comply with regulation, which in principle makes them more efficient. However, in DGE models informality introduces distortions on several margins of the economy. For example, occupational and schooling choices, the scale of operation of the firm, and the allocation of resources across plants, among others. This, in turn, affects productivity and output. In this chapter, I use a classic

[1] The views expressed in this chapter are those of the author, and do not necessarily reflect those of Banco de México.

general equilibrium model of the informal sector based on Rauch (1991) to illustrate the tradeoff between distortions originating from policies associated with the formal sector, and those associated with the informal sector. As a result of the coexistence of these two types of distortions, a U-shaped relationship between aggregate output and the size of the informal sector exists.

While early papers focused mainly on the *determinants of informality*, DGE models not only incorporate the informal sector as an endogenous object, but they also provide a framework where the determinants of informality can be linked to the performance of aggregate variables, such as unemployment and productivity. Rauch (1991) was the first to introduce informality in a general equilibrium model. He, however, was only interested in the determinants of informality and not on the effects that the policies that generate informality have on aggregates (perhaps this was due to the computational burden of the time, or the tradition of his predecessors). In this chapter, I will perform a numerical exploration of the aggregate implications of informality in Rauch's model.

To the best of my knowledge, the first contemporary article that goes beyond studying the determinants of informality is Antunes and Cavalcanti (2007). These authors introduce a model to analyze how much of the differences in the size of the informal sector and on income per-capita across countries can be accounted for by differences in the regulation of entry and on the level of enforcement of financial contracts. They find, among other results, that both determinants are important to explain the differences in informality between the U.S. and Peru. Moreover, their results suggest that the distortions induced by these two policies have a negative impact on aggregate output and Total Factor Productivity (TFP).

One challenge in this literature is to find a systematic way of identifying the determinants of informality that are relevant for aggregate outcomes. Motivated by the influential work of De Soto (1989), the regulation of entry has received substantial attention over the last two decades. Similarly, credit constraints seem to be a preferred determinant of informality by many authors (see, e.g., Amaral and Quintin, 2006, 2010; D'Erasmo and Moscoso Boedo, 2012). Nonetheless, the relevance of these two determinants has been challenged by important developments in the literature. First, regarding credit constraints the research of Midrigan and Xu (2013) and Moll (2014) shows that when constrained entrepreneurs are allowed to accumulate assets, self-financing can undo the capital misallocation that arises from financial frictions and make TFP distortions vanish in the long-run. Thus, when we try to focus on the determinants of informality that are important for aggregate outcomes in the long-run, credit constraints seem irrelevant. This is corroborated empirically. Following the discussion in Kaplan et al. (2007), the evidence suggests that informal firms are not financially constrained. In ENAMIN (2008), the owners were asked about the way they financed their start-up costs; 50.5% mentioned that they used their own savings, while only 1.2% used credit from commercial banks. This may mean that access to credit for small firms is difficult, or alternatively, that demand for credit is low. However, the owners were also asked which are the main problems faced by their business; only 2.39% answered that lack of credit was a problem. Therefore, informal entrepreneurs do not seem to be affected by credit constraints.

Second, empirical studies have found that changes in the regulation of entry have little impact on informality. In fact, Kaplan et al. (2007) suggest that the reason for this might be that even though entrepreneurs save entry costs because of the reforms they are also aware that they will have to face a large tax burden if they decide to enter the formal sector.

In my own recent work (Leal, 2010) I concentrate on the tax compliance aspect of informality. I study how the way firms avoid taxes affects output and productivity. In response to incomplete tax enforcement I find that firms reduce capital demand, which in turn impacts on capital accumulation. Incomplete tax enforcement also distorts occupational choices in the economy and misallocates resources across firms. I find that if Mexico was able to increase its tax enforcement capacity to a level high enough to eliminate informality entirely, productivity and capital accumulation would increase, which translates into an output increase of around 34%, relative to the output level in the economy with low enforcement and informality. Nonetheless, the issue of identifying the relevant determinants of informality is still an open debate.

In the next section, I continue with a brief history of the concept of informality in order to put into perspective its recent developments. Then, I discuss the main features of a class of models with heterogeneous firms and informality. I do this in the context of the pioneering work of Rauch (1991). In Section IV.4, I analyze the basic distortions induced by informality; and in Section IV.5, I discuss some variants to the basic model of informality. Finally, in Section IV.6, I provide some thoughts on the paths that the literature could take in the forthcoming years, as well as discussing the areas of opportunity for future research.

IV.2 A BRIEF HISTORY OF THE INFORMAL SECTOR CONCEPT

The concept of "the informal sector" has been present in the economics profession for more than 40 years. Its origin is attributed to Hart (1970) (see, e.g., Sethuraman, 1976) and to an International Labor Office (ILO) staff report after an employment mission to Kenya in 1972 (see, e.g., Bangasser, 2000).[2] The context in which the informality concept was born is important because it determined the path that the literature would take for the first few decades. For mainstream development economists in the 1950's and 1960's, efforts had to be put on capital formation and infrastructure; labor issues were generally of second interest. The ILO's employment missions were meant to attract the attention of development experts on employment, particularly on the deficit of *modern* jobs in developing countries and the way this deficit was being filled by *informal* income opportunities.

The development paradigm back then was grounded on the dualistic view of development (Lewis, 1954) which assumed that a worker could earn different wages depending on the sector (s)he was hired; equivalently, it assumed that formal jobs were better than informal jobs. Therefore, *subsistence labor* (i.e. non-modern / informal jobs) were simply a symptom of underdevelopment: as industrialization would take off, workers would be absorbed into the formal sector.

With time, panel data of employment surveys became widely available and an important piece of information was brought into the analysis: there exists substantial work-flows into and out of informality. This data shook the foundation of the dualistic view of informality because it provided evidence that informal jobs were not necessarily worse than formal jobs.

As a result of these historical events, a considerable amount of effort within the informal sector literature has been placed on the dualistic/segmented vs. free mobility labor markets debate. The debate still continues to this day (see Fields, 2004; and Perry et al., 2007, for two recent contrasting views). Thus, much of the literature has focused on the optimal decisions of workers, and its implications on employment levels and the size of the informal sector. Only recently the focus of researchers has started to move towards the behavior of firms. There is now a renewed interest on informality due to the availability of firm level data. An increasing number of authors are studying the decisions of informal firms, and how institutions and policy affect them. Today, the informal sector is seen as a source of resource misallocation. This has brought a twist in the informal sector literature, as many researchers now recognize the role of the informal sector as a potential cause of underdevelopment.

[2] "Employment, incomes and equity: a strategy for increasing productive employment in Kenya" (ILO, Geneva, 1972).

IV.3 A CLASS OF GENERAL EQUILIBRIUM MODELS

I will focus on a class of models in the informal sector literature that includes two main ingredients. The first ingredient is heterogenous firms. As mentioned in the previous section, informal firms tend to be small so the use of heterogenous firms in the model is the first step to deliver this equilibrium characteristic. The second ingredient consists of a group of features in the model that define the net-gain of operating informally. In this class of models, the net-gain of informality is decreasing with firm size and thus only small firms have an incentive to choose informality.

This second ingredient is one aspect where the literature takes different approaches. Typically there is a gain enjoyed by informal firms: this gain could be due to tax evasion, avoidance of entry costs, or lack of compliance of minimum wages, among other things. Additionally, there is a cost that increases with firm size. This cost can be, for example, in the form of more stringent tax/regulation enforcement for larger firms or in the form of allowing formal credit only for larger firms. This is why the net gain (gain minus cost) is decreasing with firm size.

These two ingredients are closely related to those used in the macro literature that study idiosyncratic distortions across heterogenous plants. In this literature, identified in the work of Restuccia and Rogerson (2008) and Hsieh and Klenow (2007), while firms face policy distortions that are more general than the distortions modeled by the informal sector literature, they certainly share the same spirit. Therefore, in Section IV.4 I will present some advances from this macro literature that complement the findings of the informal sector literature.

I would like to start with the first formal contribution in the informal sector literature associated with Rauch (1991). This model was influenced by the wave of evidence showing *reverse mobility*, i.e., workers moving from formality to informality. He was, at the same time, not very inclined to abandon the dualistic view of labor markets. Thus, Rauch developed a model where even by assuming free mobility across sectors a sort of segmentation occurs in the labor market. Here, I present a simplified version of Rauch's model.

IV.3.1 The distortion-free economy

The environment departs from Lucas (1978)'s occupational choice model. There is a continuum of individuals with ability $z \in [0, \bar{z}]$ and distributed according to cdf $F(z)$. Consider first the simplest version of the model without government and informality, that is, without distortions. The individuals can choose to operate either as entrepreneurs or as employees. If an individual decides to be an employee (s)he receives the wage w; on the other hand, if the individual chooses to be an entrepreneur (s)he has access to a production function with decreasing returns to scale: $y(z) = zl^\theta$, with $0 < \theta < 1$. Now consider the problem of an entrepreneur with ability z, given w:

$$\pi(z, w) = \max_l \left\{ zl^\theta - wl \right\}.$$

For later reference let $l(z, w)$ be the optimal labor demand for individual z and $y(z, w)$ be optimal output. Notice that $\pi(z, w)$ is increasing in z (an application of the envelope theorem), while w is just a constant. Not surprisingly, for any $w > 0$ there exists a threshold $\hat{z}$, such that $\pi(\hat{z}, w) = w$, to the left of which all individuals are better-off being employees (i.e. $w > \pi$), and to the right of which individuals are better-off being entrepreneurs (i.e $w < \pi$). Figure IV.1 shows how $\hat{z}$ is determined.

A competitive equilibrium for this economy consists of a threshold $\hat{z}$, wages w^*, and quantities $l(z, w)$, $\forall z \geqslant \hat{z}$, such that:

1. Each entrepreneur solves his/her problem;
2. $\pi(\hat{z}, w^*) = w^*$;
3. The labor market clears: $\int_{\hat{z}}^{\bar{z}} l(z, w^*) dF(z) = F(\hat{z})$.

By standard arguments this equilibrium exists and is unique. Lucas (1978) shows that the equilibrium allocation of this undistorted economy is also the solution to a social planner's problem that maximizes aggregate output.

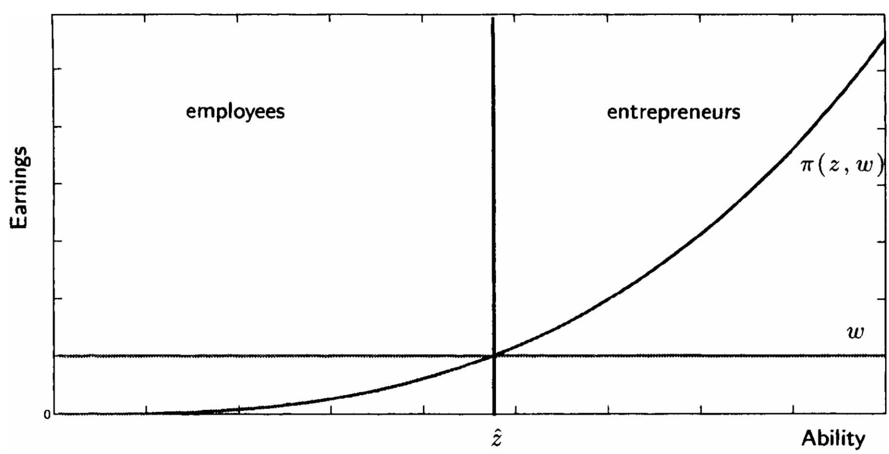


Figure IV.1. Rauch's distortion-free economy

In terms of the equilibrium characterization, it can be shown that the equilibrium threshold is given by:

$$\hat{z} = \left[\frac{\theta}{1-\theta} \frac{\int_{\hat{z}}^{\bar{z}} z^{\frac{1}{1-\theta}} g(z) dz}{G(\hat{z})} \right]^{1-\theta}, \tag{IV.1}$$

which is one equation with one unknown. The parameter θ measures the returns to scale at the firm level, which is important in determining the threshold $\hat{z}$. Thus, the returns to scale parameter is crucial to understand how resources are allocated across establishments in equilibrium. To see this, consider the problem of a social planner in the extreme case when $\theta = 1$. In this case all resources are allocated to the most productive establishment and $\hat{z} = z_{max}$. As θ moves away from one, decreasing returns become important, and allocating some resources to other firms becomes efficient.

Finally, notice that aggregate variables are given by the sum of individual decisions. For example, aggregate output is given by $Y = \int_{\hat{z}}^{\bar{z}} y(z, w^*) dF(z)$, the mass of entrepreneurs in equilibrium is given by $EntrMass = 1 - F(\hat{z})$, and the average entrepreneurial ability is: $AvEntrAb = \left(\int_{\hat{z}}^{\bar{z}} z^{\frac{1}{1-\theta}} dF(z) \right) / (1 - F(\hat{z}))$.

IV.3.2 A minimum wage economy with full enforcement

As a second step, Rauch (1991) considers an economy with perfect enforcement of minimum wages. Note that this policy is naturally distortionary: since the government does not have any other role in the model, minimum wages simply introduce a deadweight loss. The author argues that in this economy there is no equilibrium, and that labor supply will exceed labor demand when $\bar{w} > w^*$, where $\bar{w}$ represents the minimum wage and w^* denotes the equilibrium wage. Rauch (1991) concludes that in this economy unemployment will arise.

The problem of entrepreneur z is as follows:

$$\pi_{FE}(z, \bar{w}) = \max_{l} \left\{ z l^{\theta} - \bar{w} l \right\}.$$

For future reference let $l_{FE}(z, \bar{w})$ denote the optimal labor choice of entrepreneurs under full enforcement. The threshold $\hat{z}_{FE}$ is defined in a similar way as $\hat{z}$ in Section IV.3.1, except that the equilibrium wage in condition 2 above is replaced by the minimum wage: $\pi(\hat{z}_{FE}, \bar{w}) = \bar{w}$. Notice also that condition 3 above is not required. With this, we can calculate aggregate variables. However, these are not the result of an equilibrium outcome as in the case of the undistorted economy.

Distortions induced by formality. The introduction of a minimum wage increases the cost of labor to all entrepreneurs, thereby affecting production decisions. As explained above, under full enforcement there will be an excess supply of labor, which makes it difficult to compare the undistorted outcome with the distorted outcomes. Nonetheless, by inspection of Figure IV.1 it is clear that $\hat{z}_{FE} > \hat{z}$. The reason for this is that with the introduction of the minimum wage, the horizontal straight line shifts up, while the curved line shifts down and to the right. Thus, aggregate production cannot be efficient under these circumstances because a group of entrepreneurs is missing, and less labor is being used in production as a result of the increase in labor costs ($l_{FE}(z, \bar{w}) < l(z, w^*)$, $\forall z \in [\hat{z}_{FE}, z_{max}]$).

IV.3.3 A minimum wage economy with imperfect enforcement

The next step consists of introducing informality into the model. Rauch (1991) uses a cost of informality that varies with firm size, which has now become standard in the literature.[3] In his model the minimum wage ($\bar{w}$) is only enforced on firms larger than a certain size $\bar{l}$. This can alternatively be rationalized using a probability of detection equal to 0, if labor is less than or equal to $\bar{l}$, and equal to 1, otherwise (see, e.g., de Paula and Scheinkman, 2007; Bigio and Zilberman, 2011). Since more informality implies a lower compliance with minimum wages, the economy tends to move towards efficiency as informality increases. However, informality introduces other type of distortions as well, which will be analyzed in Section IV.4. In this section I will focus on describing the way informality is generated in Rauch's model.

Employees that are employed by firms that do not comply with minimum wage requirements are labeled *informal*; while employees of firms that comply with such regulation are called *formal*. Note that there is never going to be a case that an individual chooses to operate as an informal entrepreneur and at the same time chooses labor greater than $\bar{l}$, otherwise the entrepreneur will be caught and punished. Therefore, informal firms enjoy lower labor costs $w < \bar{w}$, but they also suffer a constraint on the amount of labor they can hire $l \leq \bar{l}$. Notice that this implies that the net gain of being informal varies with firm size. Firms with low z, which naturally tend to operate at a low scale, enjoy the reduced labor costs but do not bear any constraints: thus they are implicitly subsidized.[4] Firms with medium levels of ability find it attractive to distort the amount of labor hired in

[3] It was already a well-established empirical fact that informal firms are small.

[4] As in Rauch (1991), I assume that the minimum wage is binding. This assumption is one important critique to Rauch's paper, because in many developing countries with large informal sectors, minimum wages do not bind.

order to remain undetected by the government and still enjoy the benefit of paying low wages. The largest firms are better-off complying with minimum wages because their marginal productivity of labor is high and reducing labor to $\bar{l}$ signifies an important production loss that does not compensate the gain of lower labor costs.

While labor is homogenous in this model, the existence of limited minimum wage enforcement allows identical employees to earn different wages. This is in the spirit of the dualistic view of labor markets. Rauch (1991) assumes that employees are randomly assigned to each sector, so not every employee earns the formal sector wage $\bar{w}$, and unlucky individuals in the informal sector earn less than the minimum wage. This idea can be simplified by introducing stages in the timing of the decisions taken by agents in the model (this is implicit in Rauch's analysis).

Assume there are two stages and that individuals have to take their occupational choices in stage one and production decisions in stage two. Also, in stage 2 employees are sorted into the formal and informal sectors randomly. Given occupational choices of stage 1, there is no possibility to revert such decisions in the next stage. Furthermore, individuals make occupational choices facing uncertainty regarding the sector (and the wage) they will get if they choose to be employees. Nonetheless, in stage 1 individuals have some information at hand: they are aware of the market (informal) wage, of the minimum wage, and of the probability of getting a formal sector job. Therefore, agents in this model are faced with the following occupational discrete choice: a) be a formal entrepreneur; b) be an informal entrepreneur; or c) be an employee.

If an individual z decides to be a formal entrepreneur (s)he faces the minimum wage and solves:

$$\pi_F(z, \bar{w}) = \max_l \left\{ z l^\theta - \bar{w} l \right\}.$$

Let $l_F(z, \bar{w})$ denote the optimal labor choice of formal entrepreneurs. Instead, if an individual decides to be an informal entrepreneur the problem is:

$$\pi_I(z, w) = \max_l \left\{ z l^\theta - w l \right\} \quad s.t.\ l \leq \bar{l},$$

and similarly, let $l_I(z, w)$ denote the optimal labor choice of formal entrepreneurs. Finally, if the individual decides to be an employee the agent faces a lottery: with probability p the employee is assigned in the second stage to the formal sector earning $\bar{w}$, and with probability $1 - p$ goes to the informal sector earning $w < \bar{w}$. Thus, ex-ante employee earnings are given by $Ew = p\bar{w} + (1 - p)w$. Notice also that the probability of obtaining a formal sector job as an employee is an endogenous object equal to the amount of labor demanded by formal

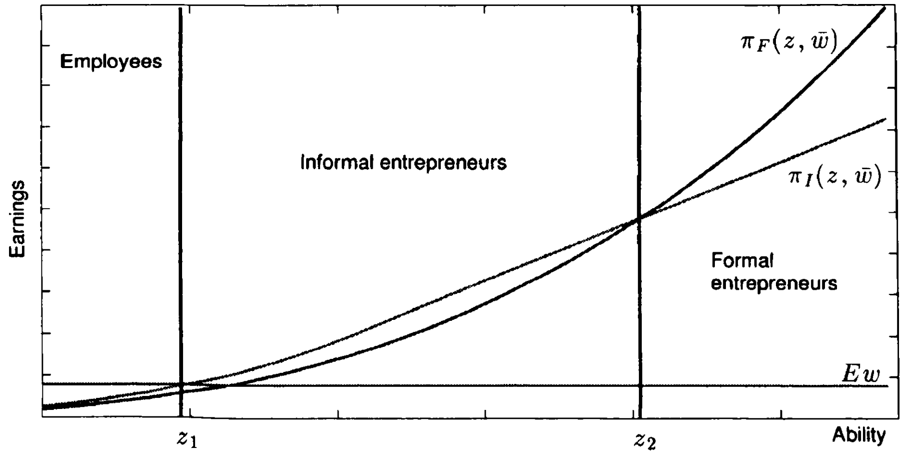


Figure IV.2. Rauch's distorted economy

firms, divided by the amount of people that are willing to become employees:

$$p = \frac{\int_{z_2}^{\bar{z}} l_F(z, \bar{w}) dF(z)}{F(z_1)}.$$

Occupational choices in the distorted economy are summarized in Figure IV.2. Entrepreneurial profits are increasing in ability and employee earnings are flat, as in the distortion-free case. Therefore, there exists a threshold z_1 to the left of which individuals are (ex-ante) better-off being employees. To the right of z_1 individuals find it more attractive to become entrepreneurs. A second feature of Figure IV.2 is that the profits of informal entrepreneurs (the $\pi_I(z, \overline{w})$ line) are higher than the profits of formal entrepreneurs for low levels of ability. This is reverted for high ability levels. The reason for this is that the constraint suffered by informal entrepreneurs becomes more costly for individuals with a large z. Given a labor level, the marginal productivity of labor is higher, the larger is z. Thus, restricting labor to $\bar{l}$ results in a larger production loss for firms with high z. This trade-off is summarized by the threshold z_2, to the left of which entrepreneurs find it more attractive to become informal, and to the right of which all of them prefer to be formal.

IV.4 Distortions induced by informality

This simple static model is rich enough to deliver interesting distortions induced by informality. There are three type of distortions: 1. on the optimal labor demand

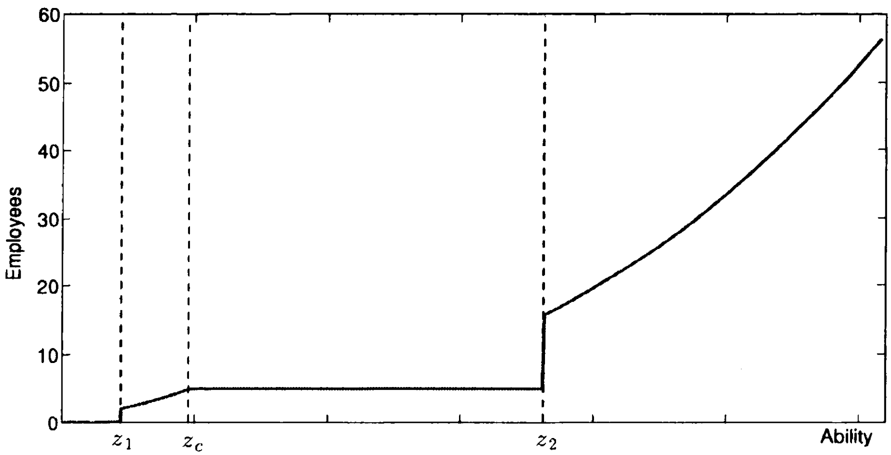


Figure IV.3. Within-firm distortion on labor choices

of informal firms; 2. on occupational choices; and 3. on the allocation of resources across firms.

The first distortion occurs within firms. It affects a group of informal firms to the left of the threshold z_2. Notice that for these firms there is an incentive to remain small in order to avoid detection. They find it optimal to operate with labor equal to $\bar{l}$ and enjoy low costs. Not every informal entrepreneur is necessarily in this situation. In particular, firms with ability just right of z_1 might not need to distort their labor choices, since they tend to operate at a low scale. This point is illustrated in Figure IV.3. Notice that between z_1 and z_c informal firms are unconstrained, while those individuals with ability between z_c and z_2 operate with labor exactly equal to $\bar{l}$. Formally, the threshold z_c is defined as the threshold that satisfies: $l(z_c; w) = \bar{l}$.

The second distortion relates to occupational choices. The limited enforcement of minimum wages makes it attractive for a group of individuals to become entrepreneurs instead of employees. This pushes the employee-entrepreneur threshold z_1 to the left, thereby distorting occupational choices. This point is illustrated by comparing Figures IV.1 and IV.2.

The third distortion is more subtle. It refers to the way existing resources are allocated across establishments, and it is intimately linked to the type of distortions found in the macroeconomic literature on resource misallocation across heterogenous plants, previously mentioned in the introduction. The main idea here is that due to the limited enforcement of minimum wages, firms with low productivity are allocated relative more resources compared to an efficient environment.

IV.4.1 A numerical example

To illustrate the mechanics of the model lets consider the following numerical example. I use a truncated Pareto distribution for entrepreneurial ability and assume that it is distributed according to:

$$G(z) = \frac{1 - \left(\frac{z_{min}}{z}\right)^s}{1 - \left(\frac{z_{min}}{z_{max}}\right)^s},$$

where $s > 0$ is called the shape parameter and $z \in [z_{min}, z_{max}]$. For this particular numerical exercise I need to provide a value for the parameters for the distribution of the entrepreneurial ability (s, z_{min}, z_{max}), the technology parameter (θ), and policy parameters: ($\bar{w}$, $\bar{l}$). I use the following values: $s = 4.25$, $z_{min} = 1$, $z_{max} = 13.05$, $\theta = 0.66$, and $\bar{w} = 0.9$. With these parameter values I perform two exercises where I change the quality of tax enforcement. First, I compute the equilibrium when $\bar{l} \to \infty$. In this equilibrium, no firm complies with the minimum wage legislation. Equivalently, this economy is operating without distortions. Second, I compute the equilibrium when $\bar{l} > 0$ is finite. In this economy, the informal sector is large and the distortions discussed earlier are present.

Table IV.1 presents the results of this exercise. When $\bar{l} = 1000$, there is no entrepreneur that complies with minimum wages (i.e. all are informal). Thus, this economy operates without distortions. When $\bar{l} = 20$, the largest firms become formal. Even though the size of the informal sector is still very large, the distortion induced by minimum wages is present and aggregate output is below the efficient level (98%). Something similar occurs when $\bar{l} = 5$. In this case the share of the formal sector is even larger and more of them are therefore forced to pay higher wages, which in turn distorts aggregate outcomes (output is reduced to 95% of the undistorted level).

Notice that both wages and the average entrepreneurial ability decrease, as $\bar{l}$ decreases. This is consistent with the employee/entrepreneur threshold z_1 moving to the left of $\hat{z}$. As z_1 moves to the left, more individuals become entrepreneurs and less employees (see the last two rows of Table IV.1). Consequently, average entrepreneurial ability goes down because the type of new entrepreneurs that enter in operation are the least productive. The mean size of establishments is also smaller, since there are less employees per entrepreneur. This result might be puzzling when compared to the result of full enforcement. In the section above we showed that threshold $\hat{z}_{FE}$ lies to the right of $\hat{z}$. The reason for this is that $\hat{z}_{FE}$ is not an equilibrium outcome, and it only has to be consistent with condition 2 above.

Table IV.1. Distortions induced by informality (numerical exercise)

	$l = 1000$	$l = 20$	$l = 5$	$l = 2.5$
Y	1.00	0.98	0.95	0.91
Share of informal employees	1.00	0.97	0.93	0.89
w	1.00	0.96	0.90	0.84
Mean size	5.91	4.93	3.77	2.72
Average entrepreneurial ability	1.00	0.96	0.92	0.86
Mass of employees	1.00	0.97	0.92	0.85
Mass o entrepreneurs	1.00	1.16	1.45	1.86

The reader might be tempted to conclude from Table IV.1, where output increases as informality increases, that the distortions associated with informality are small, and that the distortions that really matter in a quantitative sense are those associated with policies enforced in the formal sector. However, it must be stressed that this numerical exercise is not a calibration and thus it is only used for illustrative purposes. Furthermore, Rauch's model includes only one feature of informality (minimum wages) and abstracted from taxation, financial constraints, and other features that the current literature has considered important.

IV.4.2 Two sources of distortions: formality and informality

In Leal (2010), I use a model calibrated for Mexico to show that both distortions, those associated with the formal sector policies, and those that arise due to the presence of informality, are important. In fact, when the level of enforcement is plotted against equilibrium output, a U-shaped curve arises (see Figure IV.4). This means that for low levels of informality, output is high and close to the level of the undistorted economy. However, for medium levels of informality, output is low because both distortions from the formal and informal sectors are large. When informality is very low, output is again high because there are only distortions associated with formality.

IV.4.3 Idiosyncratic distortions across heterogenous plants

One particular distortion associated with informality that has been the study of recent papers identifies informality as a source of resource misallocation. This view associates informality with idiosyncratic distortions across plants. In order

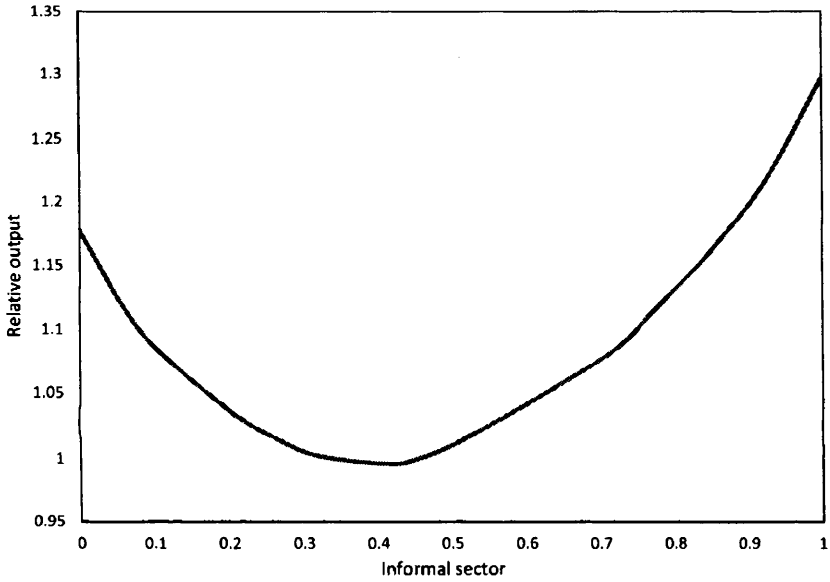


Figure IV.4. Enforcement levels (informality) and equilibrium output

to see how this can operate I quickly revise an important paper on resource misallocation and heterogenous plants by Restuccia and Rogerson (R&R, 2008). The goal of R&R is to analyze plant idiosyncratic distortions within the context of a standard neoclassical model. They depart from a simplified version of a model of firm dynamics found in Hopenhayn and Rogerson (1993).

In the R&R model, there is a representative household that lives infinitely, who does not value leisure, and has standard preferences over consumption:

$$\sum_{t=0}^{\infty} \beta^t u(C_t),$$

where C_t is consumption in period t and β is the discount factor. Endowments consists on K_0 units of capital at date 0, and one unit of time each period.

Next I briefly describe the technology. Although it is more elaborated than the previous static model, both models are designed to deliver a non-degenerated distribution of resources across heterogenous plants in equilibrium. While Lucas (1978)'s model does this using occupational choices, Hopenhayn and Rogerson (1993) develop a model of firms dynamics. Each plant has a production function $zf(k, l)$, where z is plant productivity, l is labor services, and k is capital. The function f exhibits decreasing returns to scale. Plant productivity is constant over time, but each period plants face an exogenous probability of death equal to λ. Exit realizations are iid across plants and across time. There is also an entry cost c_e that must be paid in order to create a new plant. After this cost is paid, plants take a productivity draw z from a distribution with cdf $H(z)$.

The problem faced by an incumbent plant is as follows:

$$\pi(z, \tau) = \max_{k,l} \left\{(1 - \tau)zk^{\theta_k}l^{\theta_l} - wl - rk - c_f\right\},$$

where τ is an idiosyncratic tax/subsidy that the plant bears and c_f is a fixed cost of production. Optimal input choices are given by:

$$k(z, \tau) = \left(\frac{\theta_k}{r}\right)^{\frac{1-\theta_l}{1-\theta_l-\theta_k}} \left(\frac{\theta_l}{w}\right)^{\frac{\theta_l}{1-\theta_l-\theta_k}} (z(1 - \tau))^{\frac{1}{1-\theta_l-\theta_k}};$$

$$l(z, \tau) = \left(\frac{\theta_k}{r}\right)^{\frac{\theta_k}{1-\theta_l-\theta_k}} \left(\frac{\theta_l}{w}\right)^{\frac{1-\theta_l}{1-\theta_l-\theta_k}} (z(1 - \tau))^{\frac{1}{1-\theta_l-\theta_k}}.$$

In order to see how idiosyncratic distortions misallocate resources, I will focus on one property of the equilibrium efficient allocation and describe the way this is affected by idiosyncratic distortions. For this, I take two plants z_i and z_j, such that $z_i < z_j$, and compare their corresponding optimal input choices. In the absence of distortions, relative labor depends only on the relative productivities of the two plants:

$$\frac{l(z_i, 0)}{l(z_j, 0)} = \left(\frac{z_i}{z_j}\right)^{\frac{1}{1-\theta_k-\theta_l}}.$$

In particular, this implies that in the absence of distortions, and since $z_i < z_j$, plant z_i is smaller (in terms of labor) than plant z_j. Since z_i and z_j are generic, this holds for any two plants in operation. This means that the distribution of labor is intimately linked to the distribution of z in equilibrium. However, when distortions differ across plants ($\tau_i \neq \tau_j$) relative labor is now:

$$\frac{l(z_i, \tau_i)}{l(z_j, \tau_j)} = \left(\frac{z_i(1 - \tau_i)}{z_j(1 - \tau_j)}\right)^{\frac{1}{1-\theta_k-\theta_l}},$$

which implies that the allocation of labor across plants is affected by the distribution of distortions τ across plants. To see this clearly consider the following example. Let $z_i = 2$, $z_j = 1$, and $\theta_l + \theta_k = 0.8$. Thus, the relative labor of these two plants in the absence of distortions is $\frac{l(z_i,0)}{l(z_j,0)} = 0.5^{1.25} = 0.42$. Now let one plant be subsidized and the other be taxed. Specifically, let $\tau_i = -0.5$ and $\tau_j = 0.5$. Relative labor becomes: $\frac{l(z_i,\tau_i)}{l(z_j,\tau_j)} = (3 * 0.5)^{1.25} = 1.66$. This shows that idiosyncratic distortions can potentially distort resource allocation across plants in an important amount. In this example, a plant that is half as productive as other ends up being 66% larger instead of being 60% smaller, as the efficient allocation dictates.

R&R calibrate their model to the U.S. and show that when distortions are correlated (subsidizing the least productive plants and taxing the most productive ones) total factor productivity and output can be reduced by up to 30%.

In conclusion, informality is a source of resource misallocation because it implicitly subsidizes low productive firms, while it taxes high productive firms. In addition to this, and as mentioned before, there are other distortions associated with informality. For instance, occupational choices of informal firms can also be distorted, as well as the scale of operation.

IV.5 Variants of Models with a Discrete Occupational Choice

There are four key characteristics of Figure IV.2 that define how occupational choices are made. First, employee earnings are constant across ability levels, while entrepreneurial profits are increasing. Second, notice that for low ability levels the slope of the curve of informal entrepreneurial profits is larger than the slope of formal entrepreneurial profits. Third, as z grows, the slope of the curve of informal entrepreneurial profits hits $\bar{l}^{\theta}$ and stays constant. Fourth, the slope of the curve of formal entrepreneurial profits keeps growing as z increases and eventually formal profits become larger than informal profits. These key features of Figure IV.2 (with some variants) are present in a number of general equilibrium models that focus on the informal sector (see, e.g., Gollin, 1995; Fortin et al., 1997; Choi and Thum, 2005; de Paula and Scheinkman, 2007; Quintin, 2008; Prado, 2011).

IV.5.1 Own-account workers

One of the variants the literature has considered is the inclusion of one more alternative occupation: *own-account workers* or *one-person firms* (see Gollin, 1995; 2002). The motivation of these papers stems from the fact that the labor force in developing countries is composed of a large fraction of this type of workers. The strategy to obtain a positive mass of own-account workers in equilibrium is similar to the one followed when informal firms are allowed. Typically, own-account workers cannot hire employees and only have access to the following production function: $y = Azl^{\theta}$, where z is entrepreneurial ability, $l \leq 1$ is labor, and $A > 1$ is a scale factor that captures productivity differences specific to own-account workers. Again, what is important is the slope of the own-account profits curve. For low levels of z, this slope is larger than the slope of the curve of formal entrepreneurial profits because $A > 1$. However, as z increases, the slope hits A and becomes constant. Thus, eventually as we advance on the range of abilities, employer's profits become larger than own-account profits.

IV.5.2 Capital accumulation

One more variant in the literature are models with capital accumulation (see, e.g., de Paula and Sheinkman, 2007; Leal, 2010). These models explore how informality affects not only occupational choices and resource allocation in the economy, but also investment decisions. I next describe the model used in Leal (2010).

There is a representative household populated by a continuum of individuals of mass 1 as in Guner et al. (2008). This household takes consumption decisions, accumulates capital, and makes occupational choices for each member. As in Rauch (1991), members are endowed with entrepreneurial ability z distributed in the population according to cdf $G(z)$. Each member can have one out of three alternative occupations: formal entrepreneur, informal entrepreneur, or employee. As in Rauch's paper the net gain of becoming an informal entrepreneur is decreasing with size. In this case, formal firms face an exogenous cost in the form of an output tax, where it is assumed that firms with capital larger than b face a probability of detection equal to 1. Thus, occupational choices are described by a picture similar to Figure IV.2.

The household has standard preferences:

$$\sum_{t=0}^{\infty} \beta^t u(C_t),$$

where C_t is consumption in period t and β is the discount factor. Endowments consists on K_0 units of capital at date 0, and one unit of time each period for each member. As is standard, capital accumulates according to $K_{t+1} = X_t + (1 - \delta)K_t$. The household chooses sequences for consumption, capital, and each member's occupation, taking as given the prices, taxes τ_y, a lump sum transfer T_t, and the constraint on capital of informal firms, given parameter b to maximize lifetime utility. The problem is:

$$\max_{\{C_t,K_t,I_t(z),F_t(z)\}} \left\{ \sum_{t=0}^{\infty} \beta^t u(C_t) \right\}, \tag{IV.2}$$

subject to the following budget constraint:

$$C_t(z) + K_{t+1} - (1-\delta)K_t = r_t K_t + E(w_t, r_t; \tau_y, b) + T_t, \forall t \tag{IV.3}$$

where K_0 is given, $E(w_t, r_t; \tau_y, b)$ is a function that describes earnings of each household, and $I(z; w, r)$ and $F(z; w, r)$ are index functions that represent occupational optimal decisions.

Dynamic models are useful to understand the relationship between informality and capital accumulation. In Leal (2010), the probability of being detected depends on capital and therefore capital decisions of informal firms are affected. In particular, there is a group of firms that find it optimal to operate with capital exactly equal to b in order to remain undetected by the government and still enjoy tax evasion. This feature of the equilibrium affects the aggregate demand of capital and hence capital accumulation. In a calibrated version of the model with perfect competition, I found that Mexico's output would be 17% higher if tax enforcement was improved to completely eliminate informality. A big fraction of this effect is explained by an increase in capital accumulation of 45%.

IV.5.3 Credit constraints

Amaral and Quintin (2006) build a dynamic model with credit constraints and occupational choices. As before, individuals can choose to be employees, formal entrepreneurs, or informal entrepreneurs.[5] In the formal sector profits are taxed at a rate τ, while informal managers do not pay taxes. Thus, as opposed to the model of Rauch (1991), the formal/informal tradeoff does not rely on the way taxes are enforced, but rather on the presence of financial frictions. The marginal productivity of capital is higher for individuals with high ability (ceteris paribus) and therefore they tend to operate with large amounts of capital. This makes the formal sector, where credit is available, more attractive for large firms.

The credit market works as follow. Entrepreneurs can self-finance part of their operating capital with their own savings/wealth. They can also borrow from a financial intermediary that charges an interest rate r. These loans occur within the period. There exists limited enforcement of financial contracts, and borrowers can choose to default on their debts. This is another aspect where formal and

[5] See also Antunes and Cavalcanti (2007).

informal managers differ. If a formal entrepreneur chooses to default they are subject to a penalty that represents a fraction ϕ of their income. This parameter captures the strength of contract enforcement in the economy. In contrast, it is assumed that informal firms are not subject to this penalty. Notice that this creates incentives for informal entrepreneurs to default. Since debt contracts must be self-enforcing, this also means that informal entrepreneurs will not have access to credit in equilibrium.

To illustrate the mechanics of the model, I revise the entrepreneur's problem in Amaral and Quintin (2006). We can separate the problem into two stages. First, we compute profits for any level of capital that the entrepreneur can operate with. Second, the individual chooses the amount of capital that he is going to self-finance, and the amount he is going to borrow. Consider first the problem for individual z with given capital k (before taxes):

$$\pi(z,k) = \max_{l} \left\{ z k^{\theta_k} l^{\theta_l} - wl - k(1+r) \right\},$$

where for later I define $k^*(z) = \arg\max_k \pi(z,k)$.

Now let a be the value of the assets that the individual possesses. Let s be the amount of capital that is self-financed, and let d be the amount that is being borrowed from the financial intermediary. The income of a formal entrepreneur with ability z and wealth a is given by:

$$V_F(z,a) = \max_{0 \leq s \leq a,\, d \geq 0} (1-\tau)\pi(z, s+d),$$

subject to:

$$\begin{aligned}(1-\tau)\pi(z,s+d) &+ a(1+r) \\ &\geq (1-\phi)(1-\tau)\left[\pi(z,s+d;w) + (s+d)(1+r)\right] + (a-s)(1+r).\end{aligned}$$

This last restriction is the incentive compatibility constraint that states that the value of repaying the loan has to be higher than, or equal to, the value of defaulting. In other words, the bank will only lend an amount $d \geq 0$ so that defaulting is not attractive.

Amaral and Quintin (2006) show that in equilibrium the amount that formal entrepreneurs can borrow rises with their own savings and with their ability. Both results are intuitive. Given z, as savings increase, so does self-financing, which in turn, raises the opportunity cost of defaulting. Similarly, given a, as entrepreneurial ability increases, the incentive compatibility constraint become *less binding*, which increases the opportunity cost of default. The paper also shows that, given ability z, wealthier entrepreneurs are more likely to opt for the informal sector. The reason for this is that entrepreneurs can only borrow in the formal sector. Poor entrepreneurs prefer the formal sector, despite the taxes, because this way they can borrow and increase capital. However, rich entrepreneurs are able to self-finance and thus they prefer the informal sector, since $\tau > 0$. Furthermore, Amaral and Quintin (2006) show that informal entrepreneurs operate with less capital than formal entrepreneurs.

In their model, Antunes and Cavalcanti (2007) additionally include entry costs into the formal sector. They calibrate the model to the U.S. and perform a number of counterfactual exercises. For example, they show that when the parameter that controls the quality of enforcement of financial contracts decreases by a factor of 2 from the baseline economy, the size of the informal sector increase 19% and measured output decreases by 33%. In contrast, they show that endogenous variables are less responsive to changes in entry costs. For example, when the entry cost increases by a factor of 8, the informal sector size increases by only 3.5% and the official output is only 95% of its original level.

IV.5.4 Heterogenous labor services

Amaral and Quintin (2006) use a production function at the plant level that exhibits decreasing returns to scale and capital-skill complementarities. An entrepreneur with ability z operates the following production function:

$$zF(k, l_u, l_s) = z[\min(k, l_s)]^{\alpha} l_u^{\theta},$$

where $0 < \alpha + \theta < 1$, k is capital, l_u is unskilled labor, and l_s is skilled labor. Since credit is only available for formal firms, they tend to operate with more capital than informal firms in equilibrium. This also implies that formal firms rely more intensively on skilled labor because unskilled labor is a better substitute for capital than skilled labor (i.e. there are capital-skill complementarities).

Unfortunately, Amaral and Quintin (2006) did not include a quantitative assessment of the implications of their model. However, D'Erasmo and Moscoso Boedo (2012) construct a model along the same lines. They calibrate the model to the U.S. and perform a number of counterfactuals where they analyze joint changes in the costs of formality. They find a complementary effect of entry costs and financial frictions when changing the formality/informality tradeoff from the one in the U.S. to those in developing countries. Relative to the U.S., their model generates up to a 37% decrease in TFP, 10 times larger informal sectors, and a 60% decline in the stock of skilled workers when moving to parameters that characterize the economies of developing countries.

IV.6 Conclusions

This chapter has presented a review of articles that belong to a class of models in the contemporary informal sector literature. This literature is closely related to a macro literature that studies the role of idiosyncratic distortions across plants. In these models it is clear that the informal sector plays a distortionary role in the economy, which challenges the romantic view of the informal sector as a receptor of the most pure and efficient kind of economic activity, and as an engine for development. This literature emphasizes distortions associated with the presence of informality not only within firms, but across firms. This literature shares two key ingredients: heterogenous firms and a net gain of informality that decreases with firm size. In this class of models, informality distorts the scale of firms and the allocation of resources across firms.

Additionally, the literature has explored distortions on other margins, such as occupational choices, and physical and human capital accumulation. When the results of all these papers are put together, it is clear that the informal sector is not only a symptom of underdevelopment, but is also represents a cause of underdevelopment. There is, nonetheless, a tradeoff when reducing informality: less distortions vs. more taxes.

In recent years the macro literature on misallocation across plants has focused on the role played by complementarities and links among several sectors in the economy. To the extent that the informal sector is concentrated on sectors that are linked to the rest of the economy in specific ways, this is an interesting area of future research (see, e.g., Jones, 2011).

Another area of opportunity relates to the role of financial constraints on informality. The macro literature has also made recent important developments in this area. In particular, the paper of Moll (2014) asks the question of how relevant financial constraints are for resource misallocation.

References

Amaral, P., Quintin, E. (2006). A Competitive Model of the Informal Sector. *Journal of Monetary Economics* 53, 1541–1553.

Amaral, P., Quintin, E. (2010). Limited Enforcement, Financial Intermediation, and Economic Development: A Quantitative Assessment. *International Economic Review* 51, 785–811.

Antunes, A., Cavalcanti, T. (2007). Start Up Costs, Limited Enforcement, and the Hidden Economy. *European Economic Review* 51, 203–224.

Bangasser, P. E. (2000). The ILO and the Informal Sector: An Institutional History. ILO Working Paper No. 2000/9. International Labour Organization.

Bigio, S., Zilberman, E. (2011). Optimal Self-Employment Income Tax Enforcement. *Journal of Public Economics* 95, 1021–1035.

Choi, J. P., Thum, M. (2005). Corruption and the Shadow Economy. *International Economic Review* 46, 817–836.

de Paula, A., Scheinkman, J. A. (2007). The Informal Sector, Third Version. PIER Working Paper Archive.

De Soto, H. (1989). *The Other Path: The Invisible Revolution in the Third World.* IB Tausis and Co. Ltd.

D'Erasmo, P., Moscoso Boedo, H. (2012). Financial Structure, Informality and Development. *Journal of Monetary Economics* 59, 286–302.

Fields, G. S. (2004). Dualism in the Labor Market: A Perspective on the Lewis Model after Half a Century. *The Manchester School* 72, 724–735.

Fortin B., Marceau, N., Savard, L. (1997). Taxation, Wage Controls and the Informal Sector. *Journal of Public Economics* 66, 293–312.

Gollin, D. (1995). Do Taxes on Large Firms Impede Growth? Evidence from Ghana. Bulletins 7488, University of Minnesota, Economic Development Center.

Gollin, D. (2002). Getting Income Shares Right. *Journal of Political Economy* 110, 458–474.

Guner, N., Ventura, G., Xu, Y. (2008). Macroeconomic Implications of Size Dependent Policies. *Review of Economic Dynamics* 11, 721–744.

Hart, K. (1970). Small-Scale Entrepreneurs in Ghana and Development Planning. *Journal of Development Studies* 6, 104–120.

Hopenhayn, H., Rogerson, R. (1993). Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy* 101, 915–938.

Hsieh, C. T., Klenow, P. J. (2007). Misallocation and Manufacturing TFP in China and India. NBER Working Paper No. 13290, National Bureau of Economic Research.

Jones, C. (2011). Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics* 3, 1–28.

Kaplan, D., Piedra, E., Seira, E. (2007). Entry Regulation and Business Start-Ups: Evidence from Mexico. World Bank Policy Research Working Paper.

La Porta, R., Shleifer, A. (2008). The Unofficial Economy and Economic Development. SSRN eLibrary.

Leal, J. C. (2010). Tax Collection, the Informal Sector, and Productivity. Manuscript.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School* 22, 139–191.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics* 9, 508–523.

Midrigan, V., Xu, D. Y. (2013). Finance and Misallocation: Evidence From Plant-level Data. Forthcoming, *American Economic Review*.

Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? Forthcoming, *American Economic Review*.

Perry, G., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D. (2007). Informality: Exit and Exclusion. World Bank Publications.

Prado, M. (2011). Government Policy in the Formal and Informal Sectors. *European Economic Review* 55, 1120–1136.

Quintin, E. (2008). Contract Enforcement and the Size of the Informal Economy. *Economic Theory* 37, 395–416.

Rauch, J. E. (1991). Modeling the Informal Sector Formally. *Journal of Development Economics* 35, 33–47.

Restuccia, D., Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics* 11, 707–720.

Sethuraman, S. V. (1976). Urban Informal Sector: Concept, Measurement and Policy. *The International Labor Review* 114.

V. INVESTIGATING THE ZERO LOWER BOUND ON THE NOMINAL INTEREST RATE UNDER FINANCIAL INSTABILITY

JULIO A. CARRILLO[1]
Banco de México

CÉLINE POILLY
University of Lausanne

V.1 INTRODUCTION

The recent global financial crisis has prompted macroeconomists to acknowledge the risks that *imperfect* financial markets pose for the real economy.[2] In particular, aggregate fluctuations scholars now seek to understand how financial shocks propagate to the real economy and what can be done to moderate a financial-melting tsunami. Within a general-equilibrium framework, considerable work on this domain has been recently done.[3] However, most studies omit one major aspect of the recent crisis, namely that the short-run nominal interest rate in major advanced economies has reached record low levels since the wake of the crisis.[4] For example, the U.S., U.K., and the Euro Area have all endured near-zero rates since 2009. The experience of Japan is even longer. The present study contributes to fill this gap, by exploring the propagation of different types of financial shocks in the presence of near-zero interest rates.

In this chapter we examine the effects of several financial shocks on the real economy, while imposing a zero lower bound (ZLB, hereafter) constraint on the

[1] The views expressed in this chapter are those of the authors, and do not necessarily reflect those of Banco de México.

[2] In general equilibrium, a financial sector affects the real economy when financial intermediation is imperfect. There are different ways to introduce frictions in the credit market. For instance, Kiyotaki and Moore (1997) and Bernanke et al. (1999) assume asymmetric information and borrowing constraints between banks and firms. Dib (2010) and Gerali et al. (2010) assume frictions among banks. Gertler and Kiyotaki (2010), Gertler and Karadi (2011), and Christiano and Ikeda (2013) include agency costs between depositors and banks.

[3] See, e.g., Nolan and Theonissen (2009), Gilchrist and Zakrajšek (2011), Jermann and Quadrini (2012), Christiano et al., (2014).

[4] An exception is De Fiore and Tristani (2012), who investigate the optimal response of (conventional and unconventional) monetary policy to a financial shock in the presence of a zero lower bound constraint. However, they do not compare the propagation dynamics of different types of financial shocks.

nominal interest rate. Our framework builds on the traditional New-Keynesian modeling framework by including state-of-the-art real and nominal frictions. The model is upgraded with the financial accelerator of Bernanke et al. (1999), where frictions appear between the financial intermediary and the borrowing productive sector. We include in our analysis three financial shocks: an equity or *net-worth shock* that destroys the value of borrowers' collateral; a *risk shock* that increases the uncertainty of borrowers' venture projects; and a *credit-spread shock* that increases the cost of credit independently of the financial health of borrowers. The first two shocks originate from the demand side of the credit market, since they are directly related to borrowers' conditions. The third is a supply-side shock that restricts credit regardless of borrowers' characteristics. It can be thought of as a reduced-form shock that summarizes problems within the financial sector itself.

We perform our analysis in three steps. First, we study whether the propagation dynamics of each of the three financial disturbances are strong enough to push the nominal interest rate towards its zero-floor. We thus compute the probability that the ZLB constraint binds when the economy is hit by one particular shock. Second, we compare the effects of the three financial shocks under two monetary policy regimes: in "normal"times, and when the ZLB binds. Third, we explore an unconventional policy that aims to stimulate an economy immersed in a liquidity trap, which has been induced by a financial disruption. In particular, we assume that the central bank promises to keep the nominal interest rate at very low levels for a period longer than recommended by its usual policy rule.[5] This policy is known as *forward guidance*, and its benefits and limitations have been addressed by Eggertsson and Woodford (2003), among others.[6]

Following Christiano et al. (2014), the model is calibrated to a typical advanced economy. Our results are the following. First, we argue that only the credit-supply shock is likely to set the nominal interest rate equal to zero. The other two shocks lack the propagation power to do so. The reason is that investment and consumption co-move after a credit-spread shock, while their directions diverge after a net-worth shock or a risk shock. Amano and Shukayev (2012) find similar results when comparing the credit-spread shock to other non-financial shocks. We extend their analysis to a financial-accelerator model, which allows us to compare

[5] An alternative policy for helping the economy would be a fiscal stimulus. Such a policy has been treated in a model with financial frictions by Carrillo and Poilly (2013).

[6] Del Negro et al. (2012) show that a typical DSGE model over-estimates the effects of forward guidance (see also Levin et al., 2010 and Carlstrom et al., 2012). This result can be explained by the excessive response of the long-term bond yield in the model to the policy announcement. Nonetheless, the authors show that forward guidance still helps the recovery when they correct for the response of the long-term interest rate.

the propagation dynamics of different financial disturbances and their associated ZLB-inducing probabilities. Second, we show that the three shocks affect differently the path of loans. The net-worth and the credit-spread shocks imply a counter-cyclical response for loans, while the risk shock predicts a cyclical response. This result is invariant to the monetary policy regime. Christiano et al. (2014) emphasize that only the risk shock implies loan dynamics that are consistent with the data, while the net-worth shock does not. We extend their analysis in two ways: first, we show that the credit-spread shock also has counter-intuitive loan dynamics; and second, we study all financial shocks in a ZLB regime. Putting these results together, we infer that a combination of various financial shocks might be necessary to explain all features of aggregate dynamics observed during the Great Recession. Finally, our last result refers to the effectiveness of forward guidance at moderating the recession. After a credit-spread shock, we show that there exists an optimal commitment period for the central bank to keep the interest rate at the ZLB for more time than prescribed by its policy rule. In our exercise, inflation and output volatilities are minimized if the central bank announces 5 quarters of extra liquidity. In contrast, keeping the interest rate low for more periods destabilizes the economy very quickly. Our results confirm Eggertsson and Woodford (2003) findings about the theoretical benefits of forward guidance, this time tested in a financial-accelerator model. However, our results suggest that the consequences of miscalculating the optimal commitment period can be very serious for economic stability.

The chapter is organized as follows. Section V.2 describes the baseline model. Section V.3 discusses the calibration and solution strategy. Section V.4 computes the ZLB-inducing probabilities for each financial shock, and compares their associated aggregate dynamics. Section V.5 presents the implementation of forward guidance. The final section concludes.

V.2 The model

Our framework is based on the workhorse New-Keynesian model enriched with frictions in the credit market (as in Bernanke et al., 1999). As such, the model features several real and nominal rigidities that close the gap between the model's predictions and the data. We assume that households have consumption habits; that prices and wages are sticky (as in Calvo, 1983) and partially indexed to past inflation; that adjustment costs are levied on investment; and that the nominal interest rate is constrained by its zero lower bound. In addition, financial frictions arise from asymmetric information between entrepreneurs (borrowers) and the financial intermediary (lender): the latter pays a monitoring cost to observe the entrepreneur's realized return, whereas borrowers observe it for free. Agency costs result in an negative relationship between the external finance premium and the value of entrepreneurs' collateral or net worth. In this environment, a recessionary shock decreases asset prices, thus reducing the value of entrepreneurs collateral. As the cost of borrowing increases, investment demand plummets, and asset prices drop again, leading to a new round of reductions in investment. The financial accelerator mechanism amplifies, indeed, the effects of shocks.

In what follows we denote by a hatted variable, $\widehat{a}_t$, its deviation from the deterministic steady state. A variable with neither a hat nor a time subscript, a, denotes its steady-state level. The complete log-linearized system is provided in the Appendix.[7]

V.2.1 Households

Preferences. The economy is inhabited by a continuum of differentiated households, indexed by $i \in [0, 1]$. A typical household selects a sequence of consumption and savings that are invested in a financial intermediary. Households differ by the specific labor type they are endowed with, which gives them monopolistic power to set their own wage. Household i's objective is to maximize her lifetime utility, given by

$$\mathrm{E}_t \sum_{t=0}^{\infty} \beta^t \left\{ \mathbb{U}(c_t - \varpi c_{t-1}) - \mathbb{V}(\ell^h_{i,t}) \right\},$$

subject to the sequence of constraints

$$c_t + d_t \leq w_{i,t}\ell^h_{i,t} + \exp(\varepsilon_{t-1})R_{t-1}\frac{d_{t-1}}{1+\pi_t} + \frac{\Upsilon_t}{P_t} + \frac{\mathcal{A}_t}{P_t} + \mathrm{div}_t, \qquad \text{(V.1)}$$

where E_t is the expectation operator conditional on the information available in period t; $\beta \in (0, 1)$ is the subjective discount factor; $\varpi \in [0, 1)$ is the habit parameter; c_t is real consumption; P_t is the price of final goods; $1 + \pi_t = P_t/P_{t-1}$ is the gross inflation rate; $w_{i,t} \equiv W_{i,t}/P_t$ and $\ell^h_{i,t}$ denote the real wage and the labor supply of type-i household at period t; d_t are real deposits held at the financial intermediary in period t that mature in period $t + 1$; R_t is the risk free interest rate set by the central bank; and finally, div_t, Υ_t, and $\mathcal{A}_t$ denote real profits redistributed by monopolistic firms, nominal net lump-sum transfers from the government, and nominal lump-sum transfers from entrepreneurs, respectively. The term ε_t denotes a shock that affects the spread between the risk free interest rate and the rate of return on private assets. The gross nominal interest rate

[7] The full derivations of the model are available upon request.

perceived in deposits is therefore $\exp(\varepsilon_t)R_t$. A positive innovation of this shock increases the return on savings, prompting households to consume less and save more. This shock is similar to the risk-premium shock introduced by Smets and Wouters (2007). We assume that it follows an autoregressive process, such as

$$\varepsilon_t = \rho_\varepsilon \varepsilon_{t-1} + \epsilon_{\varepsilon,t}, \qquad \textit{(Credit-spread shock)}$$

where $\rho_\varepsilon \in (0,1)$ and $\epsilon_{\varepsilon,t} \sim iid(0,\sigma_\varepsilon)$. The log-linearized first-order conditions with respect to c_t and d_t are

$$(1-\beta\varpi)\sigma\widehat{\lambda}_t = \beta b \mathrm{E}_t\{\widehat{c}_{t+1}\} - (1+\beta\varpi^2)\widehat{c}_t + b\widehat{c}_{t-1}, \tag{V.2}$$

$$\widehat{\lambda}_t = \mathrm{E}_t\left\{\widehat{\lambda}_{t+1} + \widehat{R}_t - \widehat{\pi}_{t+1} + \varepsilon_t\right\}, \tag{V.3}$$

where $\sigma \equiv -\mathbb{U}_{cc}c/\mathbb{U}_c$ is the inverse of the intertemporal elasticity of substitution and λ_t is the Lagrangian multiplier on the budget constraint. Equation (V.2) defines the marginal utility of consumption. Equation (V.3) is the Euler equation, which states that the marginal sacrifice of a consumption unit must equal the marginal benefit of consuming this unit plus a compensation driven by the real interest rate on savings.

Wage setting. A typical household i acts as a monopoly supplier of type-i labor. Following Erceg et al. (2000), a competitive labor intermediary aggregates the set of differentiated labor inputs into a single labor input ℓ_t^h using a Dixit-Stiglitz technology

$$\ell_t^h = \left(\int_0^1 \left[\ell_{i,t}^h\right]^{(\theta_w-1)/\theta_w} di\right)^{\theta_w/(\theta_w-1)}, \tag{V.4}$$

where $\theta_w > 1$ is the elasticity of substitution between any two labor types. The associated aggregate nominal wage obeys $W_t = \left(\int_0^1 W_{i,t}^{1-\theta_w} di\right)^{1/(1-\theta_w)}$. Following Calvo (1983), it is assumed that at each point in time only a fraction $1-\alpha_w$ of the households re-optimize their nominal wage. The remaining households simply index their wages to past inflation at rate $\gamma_w \in (0,1)$. Since the household is a monopoly supplier, it internalizes the demand for its labor when setting its wage. The following log-linearized wage setting equation can be derived:

$$\widehat{\pi}_t^w - \gamma_w\widehat{\pi}_{t-1} = \frac{(1-\alpha_w)(1-\beta\alpha_w)}{\alpha_w(1+\omega_w\theta_w)}\left[\omega_w\widehat{\ell}_t^h - \widehat{\lambda}_t - \widehat{w}_t\right] + \beta\mathrm{E}_t\{\widehat{\pi}_{t+1}^w - \gamma_w\widehat{\pi}_t\}, \tag{V.5}$$

where $\omega_w^{-1} \equiv \mathbb{V}_\ell/\ell^h\mathbb{V}_{\ell\ell}$ denotes the Frisch elasticity of labor supply and the gross wage inflation is defined by $\widehat{\pi}_t^w \equiv \widehat{w}_t - \widehat{w}_{t-1} + \widehat{\pi}_t$.

V.2.2 Entrepreneurs

Consider a continuum, mass one, of risk neutral entrepreneurs, indexed by $e \in [0, 1]$. At the end of period t, type-e entrepreneur buys the capital stock $k_{e,t}$ at price Q_t. The entrepreneur finances her capital purchases with own internal funds and a loan borrowed from the risk neutral financial intermediary.

Capital returns. At time $t+1$, entrepreneurs rent their capital $k_{e,t}$ to intermediate firms at the real rental rate z_{t+1}. After production, they sell the (depreciated) capital to a capital producer at price Q_{t+1}. Entrepreneurs are also requested to pay a fee cap_{t+1} to the capital producer, which is intended to cover capital adjustment costs. The *undistorted* gross nominal rate of returns on capital, R^k_{t+1}, equals[8]

$$R^k_{t+1} = \frac{P_{t+1} z_{t+1} + (1 - \delta - \text{cap}_{t+1}) Q_{t+1}}{Q_t}, \tag{V.6}$$

where $\text{cap}_t \equiv \frac{\vartheta}{2}\left(\frac{i_{t+1}}{k_t} - \delta\right)^2 - \vartheta\left(\frac{i_{t+1}}{k_t} - \delta\right)\frac{i_{t+1}}{k_t}$, with $\vartheta > 0$.[9] Capital returns perceived by entrepreneurs, $\tilde{R}^k_t$, are also distorted by the credit-spread shock, ε_t, so $\tilde{R}^k_t = R^k_t \exp\left(-\varepsilon_{t-1}\right)$. A positive innovation of ε_t reduces the value of capital and shifts downwards investment demand. Along with its effects on savings, an increase of ε_t induces a co-movement between investment and consumption (see equation (V.3)).[10]

The optimal financial contract. At the end of time t, type-e entrepreneur acquires (nominal) debt equal to $B_{e,t} = Q_t k_{e,t} - N_{e,t}$, where $N_{e,t}$ is its (nominal) net worth. As in Townsend (1979) and Bernanke et al. (1999), it is assumed that every single entrepreneur's investing project is subject to an idiosyncratic shock, $\omega_{e,t+1}$. The shock is a random variable distributed log normal with mean 1 and variance $\sigma^2_{\omega,t} > 0$.[11] Both the entrepreneur and the lender do not observe $\omega_{e,t+1}$ when they

[8] For the sake of exposition, some dynamic equations are presented in their non-linear form.

[9] For further details, see the capital producer problem in the online appendix of Carrillo and Poilly (2013).

[10] In the frictionless-financial-market model of Smets and Wouters (2007), households hold both deposits and capital goods. Therefore, a no-arbitrage condition naturally creates a spread between the risk free rate and the rate of return on capital. In a model with financial frictions, there is a separation between deposit holders (households) and capital owners (entrepreneurs), so the no-arbitrage condition no longer holds. Consequently, we need to include the risk premium shock in both the household and entrepreneur problems (see equation (V.6)).

[11] This shock is an i.i.d. random variable across time and types, with a continuous and once-differentiable c.d.f., $F(\omega)$, over a non-negative support.

sign the loan contract. As in Christiano et al. (2014), we allow the variance $\sigma^2_{\omega,t}$ to vary over time, which reflects changes in the risk of default of entrepreneurs projects. An increase in $\sigma^2_{\omega,t}$ widens the distribution of $\omega_{e,t+1}$, which makes the likelihood of success more uncertain. We assume the shock follows the process

$$\log(\sigma_{\omega,t}) = \rho_\sigma \log(\sigma_{\omega,t-1}) + (1-\rho_\sigma)\log(\sigma_\omega) + \epsilon_{\sigma,t}, \qquad (\textit{Risk shock})$$

where $\rho_\sigma \in (0,1)$, $\sigma_\omega > 0$, and $\epsilon_{\sigma,t} \sim iid(0,\sigma_\sigma)$. In time $t+1$, and given a threshold value $\bar{\omega}_{e,t+1}$, an entrepreneur repays its (real) debt, $b_{e,t}$, at the gross rate $r^L_{e,t+1}$ if $\omega_{e,t+1} > \bar{\omega}_{e,t+1}$. Then the threshold and the loan rate are jointly defined as

$$\bar{\omega}_{e,t+1}\tilde{r}^k_{e,t+1}q_t\tilde{k}_{e,t} = r^L_{e,t+1}b_{e,t}, \qquad (\text{V.7})$$

where $\tilde{r}^k_{e,t+1}$ is the real gross rate of capital returns. In the case where $\omega_{e,t+1} < \bar{\omega}_{e,t+1}$, the entrepreneur declares bankruptcy. In such a case, the lender pays a monitoring cost to audit the entrepreneur and gets to keep all of the entrepreneur's realized returns. This truth-telling mechanism prevents borrowers from misreporting their returns to fake a bankruptcy. For simplicity, the monitoring cost is a proportion $\mu \in [0,1]$ of the realized gross payoff to the entrepreneur's capital, i.e., $\mu\omega^k_{e,t+1}\tilde{r}_{e,t+1}q_t\tilde{k}_{e,t}$, where $q_t \equiv Q_t/P_t$ is the relative price of capital.

The optimal lending contract consists in choosing k_t and $\bar{\omega}_{t+1}$ in order to maximize an entrepreneur's expected returns, subject to the participation constraint of the lender (because of symmetry, we drop the type$-e$ subindex[12]), i.e.

$$\max_{\tilde{k}_t,\bar{\omega}_{t+1}} \mathrm{E}_t\left[\left(1-\Gamma(\bar{\omega}_{t+1})\right)\tilde{r}^k_{t+1}q_t\tilde{k}_t\right], \qquad (\text{V.8})$$

subject to

$$\mathrm{E}_t\left[\left(\Gamma(\bar{\omega}_{t+1})-\mu G(\bar{\omega}_{t+1})\right)\tilde{r}^k_{t+1}q_t\tilde{k}_t\right] \geq \mathrm{E}_t\left[r_t(q_t\tilde{k}_t-n_t)\right], \qquad (\text{V.9})$$

where $\Gamma(\bar{\omega}_{t+1})$ and $\mu G(\bar{\omega}_{t+1})$ represent the expected gross share of profits going to the lender, and the expected monitoring costs, respectively.[13] Equation (V.9) states that the lender participates in the contract as long as she is assured to receive an expected loan return equal to the opportunity costs of her funds. Since it is assumed that the lender can perfectly diversify the risk associated with the loan, its relevant opportunity cost is represented by the economy real risk free rate r_t.

[12] It is assumed that entrepreneurs have a linear utility in consumption and are subject to similar aggregate shocks, implying that $\tilde{k}_{t+1} = \int \tilde{k}_{e,t+1}\mathrm{d}e$, $r^k_{e,t+1} = r^k_{t+1}$, $\bar{\omega}_{e,t+1} = \bar{\omega}_{t+1}$ $\forall e$.

[13] For further details, see Bernanke et al. (1999) and the online appendix of Carrillo and Poilly (2013).

Let $\check{r}_t = \mathrm{E}_t\{r^k_{t+1}/r_t\}$ be the expected discounted return on capital. The first-order conditions of the above problem imply that, in equilibrium, the discounted return on capital, which denotes the external finance premium, equals the marginal cost of external finance. In log-linear terms:

$$\widehat{\check{r}}_t = \chi\left[\widehat{q}_t + \widehat{\tilde{k}}_t - \widehat{n}_t\right] + \varepsilon_t, \tag{V.10}$$

where $\chi \geq 0$ is the elasticity of $\widehat{\check{r}}_t$ to a measure of leverage $\left(\frac{q_t k_t}{n_t}\right)$. All else equal, the external finance premium (or simply the credit spread or the cost of credit) increases whenever the net worth falls. The reason is that a lower collateral raises the probability of loan default and so the lender demands a higher premium to compensate this risk. This relationship is the key feature of the financial accelerator model.

Aggregate net worth. The real aggregate net worth of entrepreneurs, n_t, is composed by their real aggregate capital gains, υ_t, and their aggregated real wage w^e_t.[14] At the end of period t, n_t is given by:

$$n_t = \gamma_t \upsilon_t + w^e_t. \tag{V.11}$$

The term γ_t denotes the probability of survival of an entrepreneur in a given period. Indeed, $1 - \gamma_t$ represents the odds that an entrepreneur exits the economy and loses all of her capital gains.[15] This assumption prevents entrepreneurs accumulating enough wealth to be fully self-financed and thus keep the credit market active at all times. We assume that the parameter γ_t follows the process

$$\log(\gamma_t) = (1-\rho_\gamma)\log(\gamma) + \rho_\gamma \log(\gamma_{t-1}) + \epsilon_{\gamma,t}, \qquad \textit{(Net-worth shock)}$$

where $\rho_\gamma \in (0,1)$ and $\epsilon_{\gamma,t} \sim iid(0,\sigma_\gamma)$. Real capital gains, υ_t, equal gross revenues from capital holdings from $t-1$ to t less borrowing repayments:

$$\upsilon_t = \check{r}^k_t q_{t-1} k_{t-1}\left(1 - \mu G(\bar{\omega}_t)\right) - r_{t-1}(q_{t-1}\tilde{k}_{t-1} - n_{t-1}). \tag{V.12}$$

[14] Following Carlstrom and Fuerst (1997) and Bernanke et al. (1999), entrepreneurs participate in the general labor market by supplying one unit of labor every period. This salary helps new entrepreneurs to start up their projects with a minimum set of collateral.

[15] We assume that the birth rate of entrepreneurs equals their mortality rate in order to keep the number of entrepreneurs constant. Entrepreneurs that exit at time t, consume their residual net worth such as $c^e_t = (1-\gamma_t)\varrho\upsilon_t$, where the complementary fraction $(1-\varrho)$ is a lump-sum transfer to households.

V.2.3 Capital producers

At time t, capital producers sell to entrepreneurs the capital stock k_t, which has been built by combining investment goods, i_t, and last period's depreciated capital stock:

$$k_t = (1-\delta)\,k_{t-1} + i_t - \frac{\vartheta}{2}\left(\frac{i_t}{k_{t-1}} - \delta\right)^2 k_{t-1}, \qquad \text{(V.13)}$$

where $\vartheta > 0$ controls the size of the adjustment cost on capital accumulation. In equilibrium, the relative price of capital, q_t, is given by $\widehat{q}_t = \vartheta\delta\left[\widehat{\imath}_t - \widehat{k}_{t-1}\right]$, and the law of motion (V.13) is $\widehat{k}_t = \widehat{k}_{t-1}(1-\delta) + \delta\widehat{\imath}_t$.

V.2.4 Firms

V.2.4.1 Final goods producers

The final good, y_t, used for consumption and investment, is produced in a competitive market by combining a continuum of intermediate goods, indexed by $j \in [0, 1]$ via the Dixit-Stiglitz production function

$$y_t = \left(\int_0^1 y_{j,t}^{\frac{\theta_p - 1}{\theta_p}}\, dj\right)^{\frac{\theta_p}{\theta_p - 1}}, \qquad \text{(V.14)}$$

where $y_{j,t}$ denotes the overall demand addressed to the producer of intermediate good j and θ_p is the relative price elasticity of demand for an intermediate good producer. Profit maximization yields typical demand functions:

$$y_{j,t} = \left(\frac{P_{j,t}}{P_t}\right)^{-\theta_p} y_t, \quad \text{with } P_t = \left(\int_0^1 P_{j,t}^{1-\theta_p}\, dj\right)^{\frac{1}{1-\theta_p}}, \qquad \text{(V.15)}$$

where $P_{j,t}$ denotes the price of the intermediate good produced by firm j.

V.2.4.2 Intermediate firms

Production technology. Type-j intermediate firms produce differentiated goods by combining labor $\ell_{j,t}$ and capital $k_{j,t-1}$. Capital services are rented from entrepreneurs. Type-j firm's total labor input, $\ell_{j,t}$ is composed by household labor, $\ell^h_{j,t}$, and entrepreneurial labor, $\ell^e_{j,t}$, according to $\ell_{j,t} = [\ell^h_{j,t}]^{\Omega}[\ell^e_{j,t}]^{1-\Omega}$. Type-$j$ intermediate good is produced with the following constant return to scale technology

$$y_{j,t} = \ell_{j,t}^{1-\alpha} k_{j,t-1}^{\alpha}. \qquad \text{(V.16)}$$

Each monopolistic firm determines its capital and labor demand in order to minimize its real cost, subject to its production technology, taking w_t, w_t^e, and z_t as given. Accordingly, labor and capital demands satisfy:

$$\widehat{w}_t = \widehat{s}_t + \widehat{y}_t - \widehat{\ell}_t^h, \tag{V.17}$$

$$\widehat{w}_t^e = \widehat{s}_t + \widehat{y}_t, \tag{V.18}$$

$$\widehat{z}_t = \widehat{s}_t + \widehat{y}_t - \widehat{k}_{t-1}, \tag{V.19}$$

where $\widehat{s}_t$ is real marginal cost.

Price setting. As in Calvo (1983), we assume that each period a monopolistic firm faces a constant probability, $1-\alpha_p$, of being able to re-optimize its price. Firm j takes the demand function (V.15) into account when setting its price. Additionally, it takes into consideration the possibility that this price remains unchanged for more than one period. If the firm cannot re-optimize its price, the latter is indexed to past inflation at rate $\gamma_p \in (0, 1)$. The following New Keynesian Phillips curve can be derived:

$$\widehat{\pi}_t - \gamma_p \widehat{\pi}_{t-1} = \frac{\left(1-\alpha_p\right)\left(1-\beta\alpha_p\right)}{\alpha_p}\widehat{s}_t + \beta \mathrm{E}_t\left\{\widehat{\pi}_{t+1} - \gamma_p \widehat{\pi}_t\right\}. \tag{V.20}$$

V.2.5 *Resource constraint*

The production of the final good is allocated to investment, total private consumption by households and entrepreneurs, public spending, and to monitoring costs paid by lenders

$$y_t = i_t + c_t + c_t^e + g_t + \mu G(\bar{\omega}_t) r_t^k q_{t-1} \bar{k}_t,$$

where g_t denotes government expenditures. For simplicity, we assume that government spending is financed with lump-sum taxes.

V.2.6 *Monetary policy*

The *gross* nominal interest rate, R_t, implemented by the central bank, satisfies a zero lower bound constraint of the form

$$R_t = \max\left(1, R_t^{not}\right), \tag{V.21}$$

where the desired (notional) nominal interest rate, R_t^{not} in gross terms, obeys the rule

$$\widehat{R}_t^{not} = \rho_R \widehat{R}_{t-1}^{not} + (1-\rho_R)\left[a_\pi \widehat{\pi}_t + a_y \Delta \widehat{y}_t\right], \tag{V.22}$$

where $\rho_R \in (0,1)$ is a smoothing parameter, a_π is the elasticity of R_t^{not} to the inflation gap (i.e. the difference of inflation from its target value), and a_y is the elasticity of R_t^{not} to output growth.

V.2.7 Equilibrium

In a symmetric equilibrium, all entrepreneurs, households, and firms are identical and make the same decisions. In addition, equilibrium in the labor market yields $\int_0^1 \ell_{j,t} dj = \ell_t^h$. The symmetric equilibrium is characterized by an allocation $\{y_t, c_t, i_t, \ell_t, k_t, n_t\}$, and a sequence of price and co-state variables $\{\pi_t, r_t, r_t^k, q_t, \pi_t^w, z_t, \lambda_t, \bar{\omega}_t\}$ that satisfy the optimality conditions in each sector, the monetary policy rule, and the stochastic shocks.

V.3 Calibration and Solution Strategy

V.3.1 Calibration

The model parameters (summarized in Table V.1) are calibrated to fit a quarterly frequency. Most values from the households and production sectors are borrowed from Christiano et al. (2014), who fit their model to U.S. data. Values from the financial sector are taken from Bernanke et al. (1999). The subjective discount factor, β, is set to 0.995, which entails an annual real interest rate of 2 per cent. The Frisch parameter, ω_w^{-1}, is set to unity. The degree of habit consumption, ϖ, is set to 0.74, while the intertemporal elasticity of substitution, σ^{-1}, is set to 1.

Regarding production, the capital share in the intermediate sector, α, is set to 0.4, and the rate of capital depreciation, $\delta = 0.025$. Following Christiano et al. (2011), the capital adjustment cost, ϑ, is calibrated to 17. Concerning price setting, we assume that the elasticity of substitution between intermediate goods, $\theta_p = 6$, which implies a price mark-up of 20 per cent. Similarly, the elasticity of substitution between labor types, θ_w, is set to 21, which translates into a wage mark-up of 5 per cent. The degrees of price and wage rigidities, α_p and α_w, are set equal to 0.74 and 0.81 respectively, implying average durations between price or wage re-optimization of about one year. Price and wage indexation parameters, γ_p and γ_w, are set to 0.10 and 0.51, respectively.

In terms of monetary policy, the interest rate smoothing parameter, ρ_R, is calibrated to 0.85; the elasticity of the notional interest rate with respect to inflation, a_π, is set to 2.40; and the elasticity of the interest rate with respect to output

Table V.1. Calibrated parameters

	Preferences and technology	Value
β	Discount factor	0.995
ϖ	Degree of habit consumption	0.74
σ	Inverse of the elasticity of intertemporal substitution	1.0
ω_w	Elasticity of labor disutility	1.0
α	Elasticity of value added wrt capital	0.40
δ	Capital depreciation rate	0.025
ϑ	Capital adjustment cost	17.00
	Nominal rigidities	
θ_p	Elasticity of substitution of goods	6.00
α_p	Degree of price stickiness	0.74
γ_p	Degree of price indexation	0.10
θ_w	Elasticity of substitution of labor	21.00
α_w	Degree of wage stickiness	0.81
γ_w	Degree of wage indexation	0.51
	Fiscal and monetary policy	
g/y	Share of government expenditure in output	0.20
ρ_R	Interest rate smoothing	0.85
a_π	Elasticity of the interest rate wrt inflation	2.40
a_y	Elasticity of the interest rate wrt the output gap	0.36
	Financial accelerator mechanism	
Ω	Share of household labor in aggregate labor	0.9833
x	Steady-state ratio of capital to net worth	2.00
$\check{r}$	Steady-state risk spread	$1.02^{0.25}$
γ	Survival rate of entrepreneurs	0.9843
$\bar{\omega}$	Threshold value of idiosyncratic shock	0.4983
σ_ω	Standard error of idiosyncratic shock	0.2763
μ	Monitoring cost	0.1175

growth, a_y, is set to 0.36. These values follow once more the estimations of Christiano et al. (2014). The steady-state share of government purchases in total output is calibrated to 0.20, which roughly corresponds to the last decade historical average.

We now turn to the parameters relating to the financial sector. The share of entrepreneurial wages in terms of income is set to 0.01, implying a value of $\Omega = 0.9833$. The steady-state share of capital investment that is financed by the entrepreneur's net worth, $x = \tilde{k}/n$, is calibrated to 2, meaning that the steady-state leverage ratio amounts to 50 per cent. The steady-state external finance premium, $\check{r} = r^k/r$, is set to $1.02^{0.25}$. The annual business failure rate, $F(\bar{\omega})$, is set to 3 per cent. It is assumed that the idiosyncratic productivity shock, ω_t, has a log-normal distribution with positive support and an unconditional expectation equal to 1. These moments help to determine the steady-state survival probability of entrepreneurs, γ, which is set to 0.9843, the monitoring costs to realized payoffs ratio, μ, which amounts to 0.1175, the steady-state variance of the entrepreneurs' idiosyncratic shock, σ_ω, which is equal to 0.2763, and the steady-state idiosyncratic threshold is set to 0.4983.

Finally, the shocks are calibrated as follows. For the credit-spread shock, we have $\rho_\epsilon = 0.95$, as in Fernández-Villaverde (2010); for the net-worth and risk shocks, we borrow from Christiano et al. (2014), who set $\rho_\gamma = 0$, and $\rho_\sigma = 0.97$. The size of the shocks vary with different exercises, which is indicated in the following sections below.

V.3.2 Solution strategy

We assume that at time $t = 0$ the economy is hit by a negative financial shock, which pushes the nominal interest rate towards its zero floor. Given its non-linear nature, the ZLB constraint prevents us from using a standard solution method. We thus adopt a piecewise-linear approach, similar to Bodenstein et al. (2009) and Christiano et al. (2011). In particular, we use an extended deterministic path over the linearized model equations, where we make sure that the ZLB constraint is satisfied at all times. We then apply a shooting algorithm to determine the duration of the liquidity trap, which is endogenous.[16,17]

[16] We encourage the reader to consult Carrillo and Poilly (2013) and its related online appendix for further details.

[17] Alternative solution methods exist, such as a collocations, spline functions, or projections (see, e.g., Nakov, 2008; De Fiore and Tristani, 2012; Fernández-Villaverde et al., 2012). However, they can be very costly in terms of computational time for complex models like ours. We also disregard equilibria driven by self-fulfilling beliefs (like in Mertens and Ravn, 2011).

V.4 Financial shocks in a liquidity trap

In this section, we investigate the effects of the three financial shocks in the economy. The net-worth shock and the risk shock originate on the demand side of the credit market. They respectively affect the value of the entrepreneur's collateral and its idiosyncratic likelihood of default. In contrast, the credit-spread shock is a supply-side disturbance that restricts credit despite the current state of the economy. It can be thought of as a reduced-form shock that summarizes problems within the financial sector itself.

Our analysis proceeds as follows. We first ask how likely is it that the economy falls into a liquidity trap. We argue that only the credit-spread shock can push the central-bank interest rate towards its zero floor. A similar result is found in Amano and Shukayev (2012), although they only compare the credit-spread shock to other non-financial shocks. Next, we explore the dynamics of an economy hit by a positive credit-spread shock in the two monetary policy regimes: when the zero bound binds, and when it does not. Finally, as we are still interested in the consequences of the other two shocks, we study the partial impulse responses of the net-worth and risk shocks in the two aforementioned interest-rate regimes.

V.4.1 What financial shock causes a liquidity trap?

We answer this question by computing the probability that the ZLB binds for each of the three financial shocks considered. What interests us is the propagation dynamics of each shock. So, we normalize the standard deviation of all financial perturbations, such that a one-standard-deviation shock increases the external finance premium by 125 basis points on impact, in annual terms.[18] That is, we assume that all shocks have, on average, the same initial effect on the credit spread of the economy. Similar to Amano and Shukayev (2012), we draw 10,000 random innovations from a normal distribution and we use them to compute a similar number of series for each of the financial shocks.[19,20] Then, we count the number of series for which the nominal interest rate equals zero for at least

[18] Certainly, the 1.25% is an ad hoc choice. However, the size of this normalizing constant does not affect the qualitative direction of our simulations. On the upside, we are able to mimic the results of Amano and Shukayev (2012) for the credit-spread shock.

[19] Amano and Shukayev (2012) identify and estimate the dynamics of the credit-spread shock and include it in a model which features other aggregate shocks. Although their approach is robust, they do not include other types of financial shocks in their analysis.

[20] We have 10,000 impulse-response series for each of the shocks, each one featuring a single innovation in period 1.

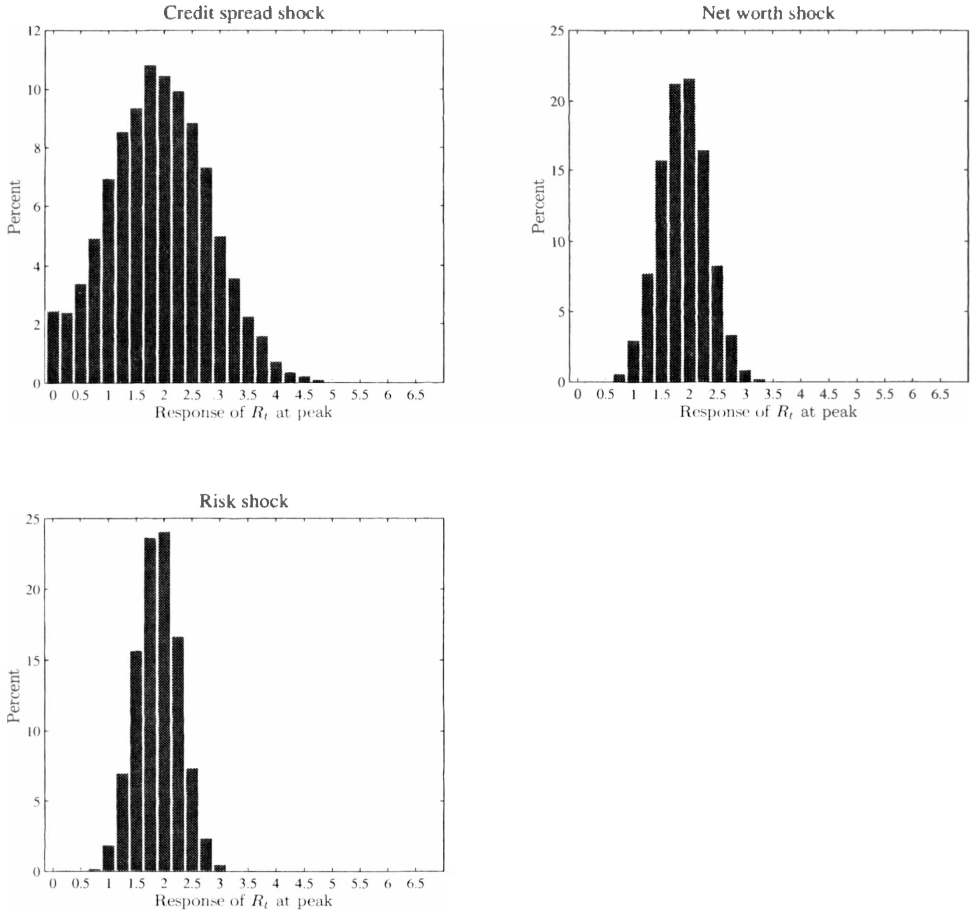


Figure V.1. Distribution of the response of the nominal interest rate at its peak effect for 10, 000 draws for the three financial shocks

one period. We obtain the following probabilities: 1.24% for the credit-spread shock, 0% for the net-worth shock, and 0% for the risk shock. Put differently, the zero bound binds on average every 20 years ($= \frac{1}{1.24\%} \div 4$ quarters) for the credit-spread shock, whereas for the other shocks it never binds. Figure V.1 displays the distribution of the nominal interest rate at the quarter where a financial shock reaches its peak propagation effect. Peak effects are reached after six quarters for the credit-spread shock and three quarters for the other two shocks. Figure V.1 shows that the nominal interest rate lies in the interval 0–0.25%, in annual terms, about 2.5% of the times for the credit-spread shock (or once every 10 years), and never for the net-worth shock and the risk shock. Amano and Shukayev (2012)

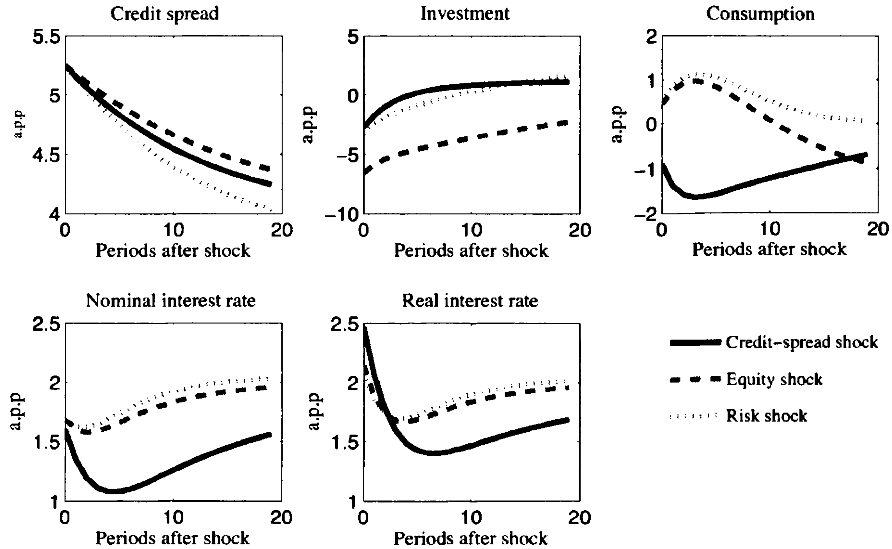


Figure V.2. Paths for investment and consumption after three different financial shocks

obtain similar findings for the credit-spread shock. For the other two shocks, we do not know of any other study that computes their ZLB-inducing probabilities.

Given the near-zero probabilities for the demand-side financial shocks, we argue that only a supply-side shock gives more chances for the nominal interest rate to hit its zero floor. The reason is that a credit-spread shock implies a co-movement between investment and consumption, whereas the other considered shocks affect mainly investment and have a counter-intuitive effect on consumption. Figure V.2 illustrates this point, where a one-standard-deviation innovation is displayed for each type of financial perturbation. Notice that the external finance premium rises from 4% in annual terms, its steady state level, towards 5.25% at impact for all the three shocks. By inspection, investment and consumption co-move only for the credit-spread shock, whereas consumption actually rises in the short-run for the other two shocks. On the one hand, a credit-spread shock achieves co-movement because it reduces the returns of capital so investment demand falls, and, on the other, it increases the returns of savings so private consumption falls. While the net-worth and risk shocks both reduce investment, they fail to discourage consumption. In fact, consumption soars because, in general equilibrium, a fall in investment demand must be met by a decrease in savings. Hence, the real interest rate must fall to discourage households to save, who use their spare resources for consumption purposes. Overall, it is more likely that the ZLB binds with a credit-spread shock as the two main components of aggregate demand fall simultaneously.

V.4.2 The dynamics of the model

In this subsection, we assume that the ZLB binds due to a credit-spread shock. First, we analyze the aggregate dynamics associated with this shock in two regimes: when the interest rate moves freely, and when the ZLB constraint is imposed. We then analyze the effects of the net-worth and risk shocks by looking at their *partial* impulse responses in each of the two interest-rate regimes. We assume that these shocks hit the economy only after the credit-spread shock has occurred. Looking at the partial impulse response functions allows us to compare the marginal effects on endogenous variables of all the three shocks.

V.4.2.1 Credit-spread shock

Let the spread between the expected returns on private assets and the risk free rate widen (i.e. a positive innovation on ε_t). This affects both the households intertemporal decisions and the expected return on capital. Gilchrist et al. (2009) and Fernández-Villaverde (2010) also assume that the external finance premium is driven by a stochastic component, and interpret it as a credit-supply shock. Figure V.3 compares the impulse response functions (IRFs, hereafter) of this shock under two cases: the non-ZLB regime and the ZLB regime. We first consider the ZLB regime. The size of the shock is deliberately chosen so that the ZLB binds at period 1 (see constraint (V.21)).[21] The monetary policy instrument stays at zero for 12 quarters, until the economy shows signs of recovery. A positive innovation of ε_t drives the external finance premium up (see equation (V.10)). As credit market conditions worsen, investment and the price of capital plummet. Despite this recessionary trend, this shock generates a counter-cyclical response of loans! Similar to Christiano et al. (2014), the intuition is as follows: the reduction in the price of capital is temporary, suggesting that the return on capital will increase in the future (see equation (V.6)). Higher expected returns on capital increases the expected value of the project, stimulating in turn the demand for loans. It is worth noting that effective loans are demand-driven in the financial-accelerator model. In fact, our perfectly competitive financial intermediary is willing to provide as many loans as entrepreneurs demand, under the condition that the latter

[21] The size of the credit spread shock is set to $\sigma_\varepsilon = 0.012$.

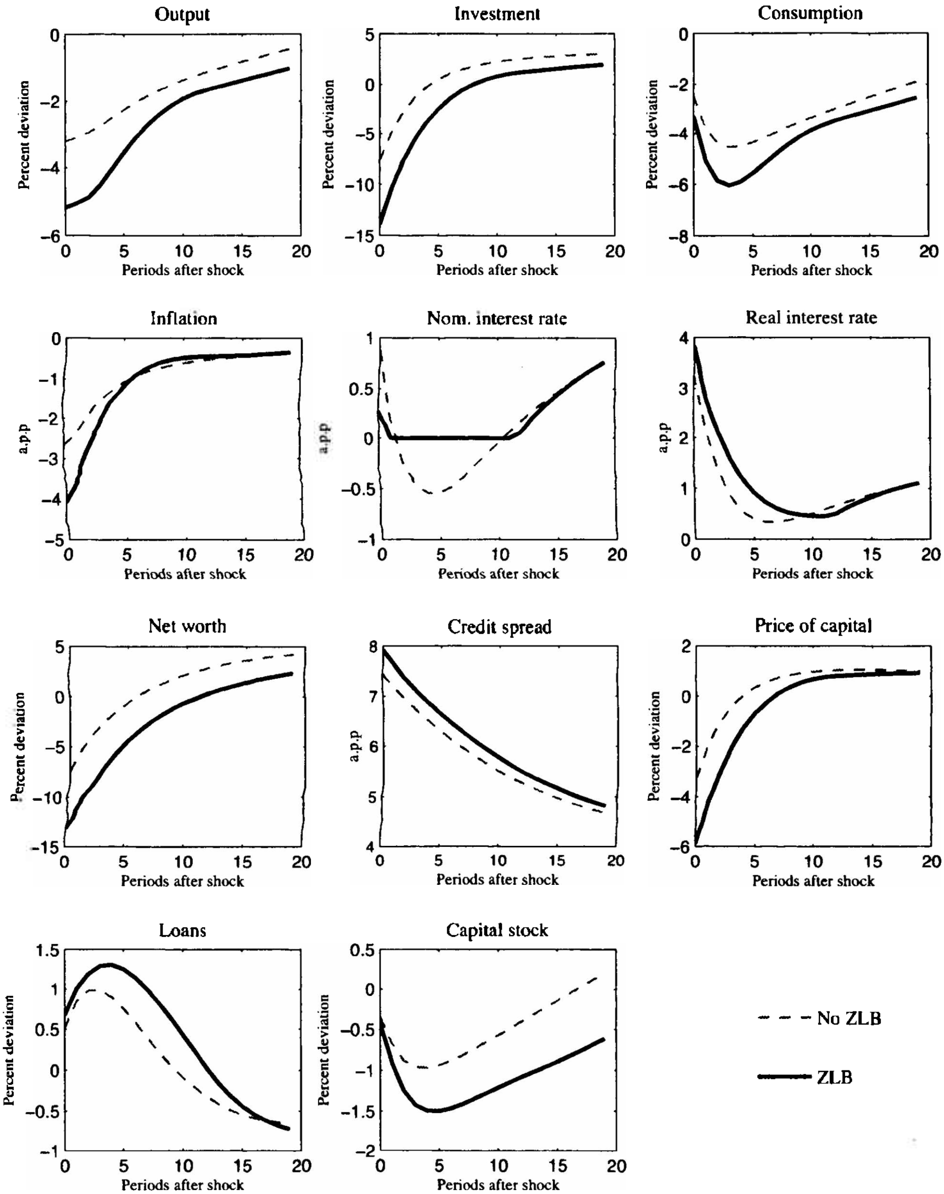


Figure V.3. Impulse responses for the credit-spread shock (selected variables) pay the current external finance premium (which ensures that the participation constraint of the lender is satisfied at all times).[22] Also, notice that investment and consumption co-move for this shock, as already noted above.

We now focus on the non-ZLB regime. As expected, Figure V.3 shows that in this case the recession is dampened. This result is not surprising, since the monetary authorities are now allowed to use their instrument to stimulate private spending and promote the recovery. While the presence of a liquidity trap makes the recovery harder to reach, it does not change the direction in the response of any endogenous variable.

V.4.2.2 Net-worth shock

We now assume a reduction in the entrepreneurs survival probability (γ_t), which can be interpreted as an exogenous decrease in the entrepreneurs net-worth value. This shock directly deals with a shift in the demand for capital through a lower aggregate purchasing power of entrepreneurs. Christiano et al. (2014) refer to it as an *equity shock*. As mentioned above, it is unlikely that the ZLB constraint binds for a reasonable size of this shock. So, in order to fix this issue, we assume that the credit-spread shock drives the nominal interest rate towards the ZLB, and that the net-worth shock occurs once the ZLB is in place. Figure V.4 compares the *partial* IRFs of selected variables in response to a negative shock on the survival probability, γ_t.[23] It is worth noting that the partial IRFs isolate the effect of the net-worth shock by showing marginal effects, which allows for the comparison with the previous exercise.[24]

In the non-ZLB regime, the model predicts that the nominal interest rate drops for several quarters. In addition, the negative financial shock reduces the demand for capital which drives its price down. Through the financial accelerator mechanism, the drop in the value of the collateral raises the external finance premium and depresses investment, leading the economy into a recession. As emphasized by Christiano et al. (2014), this shock generates a counter-cyclical demand for loans. As explained above, when the return on capital is expected to rise, the demand for loans increases. Regarding consumption, it rises in the medium-run, as in Christiano et al. (2014). The reason for the lack of co-movement between

[22] A different response of loans might be obtained in a model which explicitly describes the behavior of banks, setting some frictions to the supply of loans.

[23] Precisely, we compute the partial IRFs as the difference between the IRFs to both financial shocks and the IRFs to the net-worth shock. Since the partial IRFs are the result of a difference, they are expressed in percentage points and not in percent deviations from its steady-state level.

[24] The size of the shock is set to -0.0081, which is equal, in absolute value, to the estimated standard deviation of the equity shock in Christiano et al. (2014).

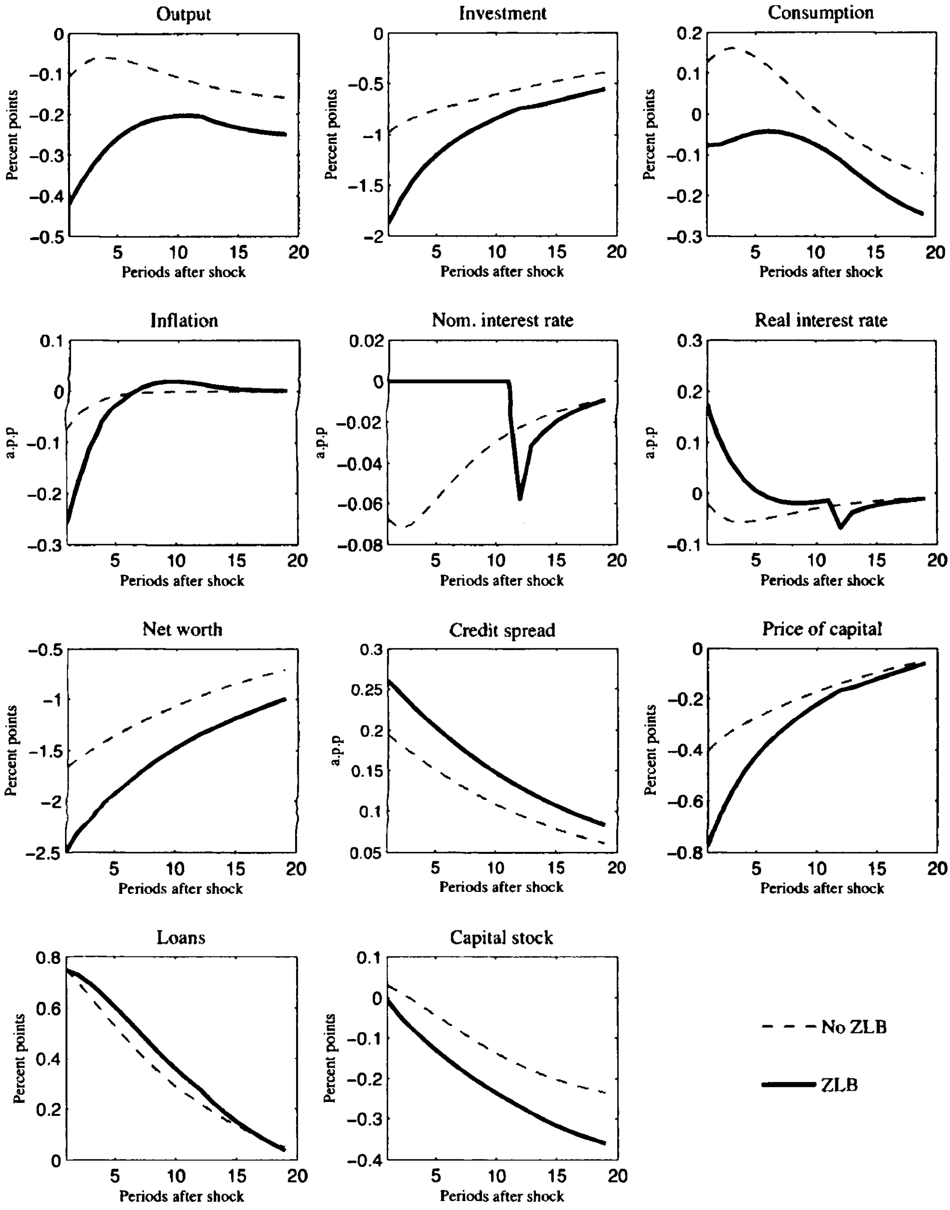


Figure V.4. Partial impulse responses for the net-worth shock (selected variables) investment and consumption for demand-side financial shocks is stated in Section V.4.1 above.

Let us now assume that the economy is in a liquidity trap (the ZLB regime). In this case, a negative net-worth shock generates more volatility. The deflationary effects of this shock generate a strong rise in the real interest rate. Consequently, the demand for capital drops by more than the price of capital. As a result, the response of loans is slightly more counter-cyclical. In addition, investment and output fall sharply, and now consumption also falls. Thus, investment and consumption co-move only during the ZLB regime. However, it has been established above that a net-worth shock on its own cannot generate a liquidity trap.

V.4.2.3 Risk shock

The last financial shock we consider is a rise in volatility of the idiosyncratic productivity shock (σ_{ω}). A risk shock widens the entrepreneurs' returns distribution and eventually worsens credit conditions (the quality of entrepreneur projects is hardly distinguishable). This shock has been analyzed by Christiano et al. (2014) in the absence of a liquidity trap. Similar to the net-worth shock, a liquidity trap cannot be reached with a risk shock. Figure V.5 compares the *partial* IRFs of this shock in the non-ZLB regime, and in the presence of a liquidity trap. Consider first the non-ZLB regime. An increase in project uncertainty makes it harder for the lender to distinguish whether an entrepreneur might default or not. The latter translate into a rise in the lending contract threshold, $\bar{\omega}_t$, and consequently, more entrepreneurs default. Therefore, the risk premium increases, and the demand for loans falls, which is now procyclical. As in the previous demand-side shock, inflation initially drops, following marginal cost and output. Interestingly, our results differ from Christiano et al. (2014) regarding the effect of the risk shock on consumption. We observe a medium-run increase in consumption when risk is rising. This difference might be explained by the presence of a variable utilization rate of capital and investment adjustment costs (rather than capital adjustment costs) in their paper.

Now consider the ZLB regime. Here, the real interest rate strongly increases due to deflationary pressures. As before, consumption is reduced, at least in the short-run. In addition, since the nominal interest rate is stuck at its zero floor, the risk shock generates a much larger increase in the credit spread, leading to a strong reduction in investment and a large recession.

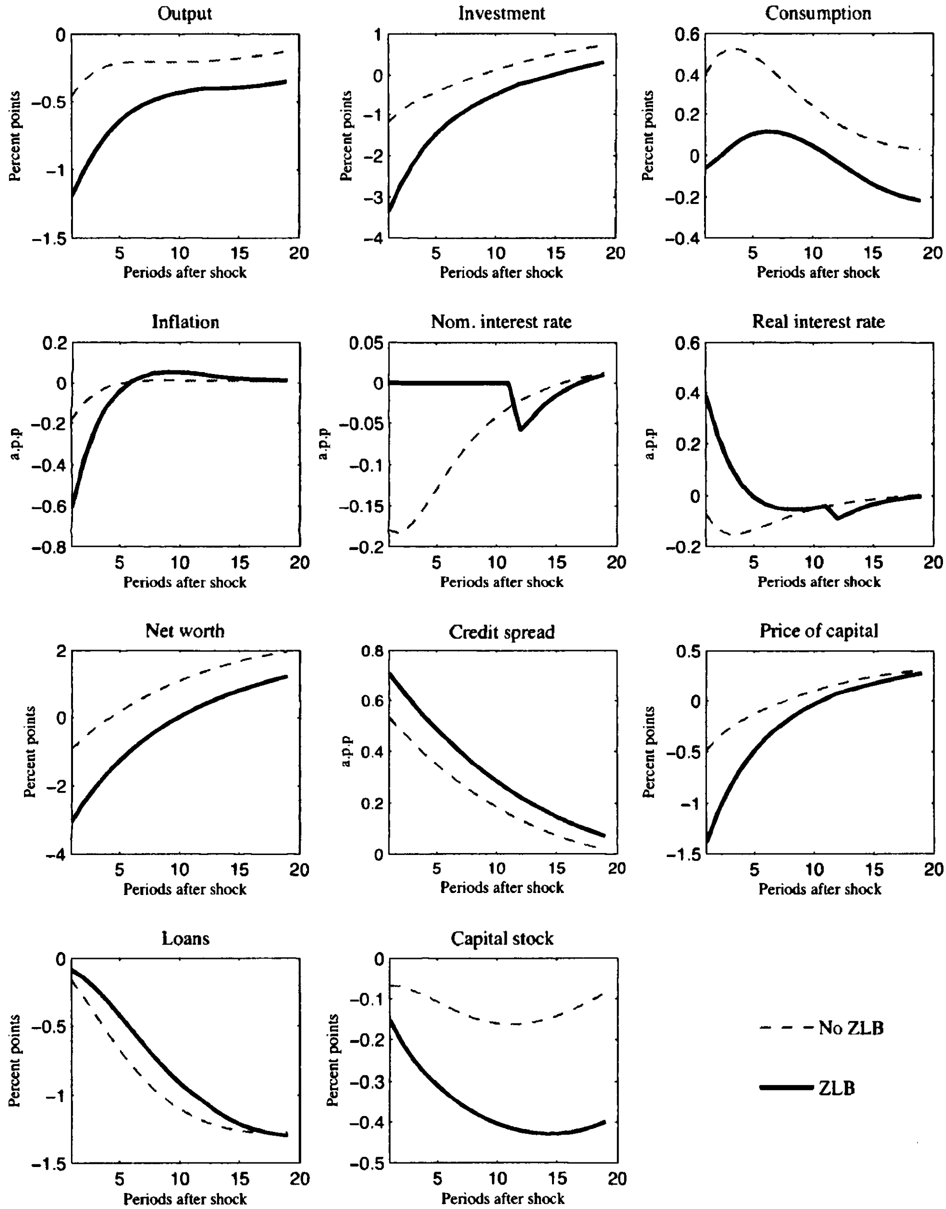


Figure V.5. Partial impulse responses for the risk shock
(selected variables)

V.5 THE CREDIT-SPREAD SHOCK AND FORWARD GUIDANCE

This section is motivated by the findings of Eggertsson and Woodford (2003). In an extensive discussion, these authors explore alternative policies that the monetary authority may employ when dealing with a liquidity trap. In particular, they explore *forward guidance*, a policy in which the central bank announces and keeps the nominal interest rate at very low levels for a longer time than prescribed by a strict inflation targeting rule. The argument, accordingly, is that when agents expect a period of abundant liquidity, they smooth consumption and plan investment better.

We focus on the credit-spread shock, since it constitutes a good candidate to generate a liquidity trap. In the initial scenario, recall that the Taylor rule prescribes that the interest rate should be positive after 13 periods. In the alternative policy, the central bank announces and keeps the nominal interest at its ZLB for more quarters than recommended by its Taylor rule. Our aim here is to investigate how inflation and output volatility react to this news. We thus perform an heuristic evaluation of this policy using a representative loss function that may represent the preferences of the central bank of the form:[25]

$$\sum_{t=0}^{\infty} \beta^t \left(\hat{\pi}_t^2 + \lambda (\hat{y}_t - \hat{y}_t^f)^2 \right). \tag{V.23}$$

In this numerical exercise, we assume that deviations of output are as important as the deviations of inflation from its target value, thus $\lambda = 1$. Notice as well that the output gap is defined with respect to the level of output that would prevail in the absence of nominal rigidities, $\hat{y}_t^f$. Finally, we compute the loss function at convergence (we set the end period 300 quarters after the shock). The upper panels of Figure V.6 display the responses of output and inflation gaps to the financial shock when the ZLB constraint is binding for 13, 18, and 21 quarters. The lower panel of Figure V.6 provides the value of the loss function (V.23) for different durations of the liquidity trap. It clearly appears that the loss function is convex with respect to the number of periods the constraint is binding. This convexity suggests that there exists an optimal commitment period for the central bank to keep the interest rate at low levels. The welfare losses are minimized

[25] This standard loss function is commonly used in the literature, which can be derived from a second-order Taylor expansion from the utility function of the representative household. In the derivation of the optimal monetary policy, Oda and Takashi (2008) interpret the managing expectations policy as the "zero interest rate commitment". We adopt here a more elementary approach since we are not interested in issues of optimal monetary policy.

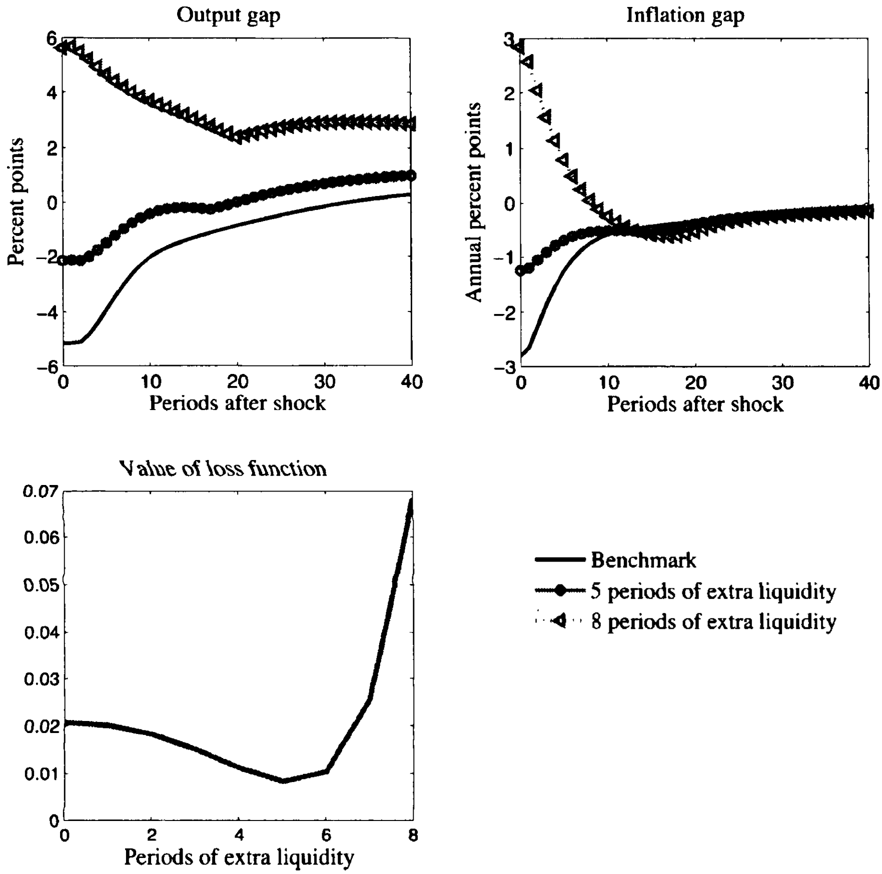


Figure V.6. Forward guidance under a credit-spread shock

when the central bank commits to keeping low the interest rate for 18 periods, i.e. 5 quarters more than prescribed by its Taylor rule. Beyond that point, the value of the loss function increases, implying that inflation and output become too volatile.

This last result is illustrated in the upper panel of Figure V.6. Precisely, the commitment period alters significantly the response of real variables. Lets consider the optimal commitment period (5 quarters). Since agents expect a low interest rate for a long period of time, they respond to the adverse economic conditions by smoothing their consumption and investment. This yields a lower deflation, which in turn produces a smaller increase in the real interest rate. On the one hand, a smaller deflation rises by less firm's real debt through the debt-deflation channel

(Fisher, 1933). On the other hand, a lower increase in the real interest rate also reduces the service of the debt. These two effects, which can be observed in equation (V.12), stop the net worth from falling deeply. Although the economy still experiences a recession, the volatilities are smaller than in the benchmark case. When the commitment period is larger than 5 quarters monetary policy avoids the recession, but at the cost of stronger volatility in the inflation and output gaps. The bottom line of this analysis is that there exists an optimal commitment period that the central bank must consider in order to provide abundant liquidity, which is in line with the analysis of Eggertsson and Woodford (2003). Further than this period, the economy destabilizes very quickly.

V.6 Conclusions

This chapter studies the propagation of different types of financial shocks in the economy, given the presence of the ZLB constraint on the nominal interest rate. In a financial-accelerator model, we look at a net-worth shock, a risk shock, and a credit-spread shock. The first two shocks originate from the demand side of the credit market, while the third one is a supply-side shock that creates a gap between the interest rate of private assets and the risk free rate. We first ask which one of these shocks is strong enough to bind the ZLB. We find that only the credit-supply shock is likely to do so, as the other two shocks lack the necessary transmission channels. The credit-spread shock causes investment and consumption to move in similar directions, whereas for the other shocks investment falls but consumption booms. We also find that, out of the three shocks, only the risk shock implies a procyclical response of loans, which is typical of a recession. For the other two shocks, loans are counter-cyclical, which is at odds with the data. Finally, we show that, after a negative financial shock, there exists an optimal commitment period for the central bank to keep the interest rate at the ZLB for more time than prescribed by its policy rule. However, the volatility of inflation and output rise quickly and sharply after this optimal period. Thus, forward guidance may be destabilizing when not used properly.

As it is widely known, the vulnerability of an ill-behaved banking sector was a crucial determinant of the global financial crisis. In our model, we have focused on the traditional financial accelerator model to study the transmission of shocks, while assuming that the financial intermediary is fully insured and perfectly competitive. Given the potential of the credit-spread shock to generate a liquidity trap, a natural step in the analysis of the propagation of financial shocks is to include an imperfect banking sector in which the credit-spread shock is micro-founded.

This modification will prove very valuable for the evaluation of anti-recessionary policies, and it might help to obtain a response of loans more in accordance with the data. We leave this topic for future research.

V.7 Appendix

The complete log-linearized model is summarized by the following equations:

$$(1-\beta b)\sigma\widehat{\lambda}_t = \beta b \mathrm{E}_t\{\widehat{c}_{t+1}\} - (1+\beta b^2)\widehat{c}_t + b\widehat{c}_{t-1}, \tag{A.1}$$

$$\widehat{\lambda}_t - \widehat{R}_t - \varepsilon_t = \mathrm{E}_t\left\{\widehat{\lambda}_{t+1} - \widehat{\pi}_{t+1}\right\}, \tag{A.2}$$

$$\widehat{r}_t = \widehat{R}_t - \mathrm{E}_t\left\{\widehat{\pi}_{t+1}\right\}, \tag{A.3}$$

$$\widehat{\pi}^w_t - \gamma_w\widehat{\pi}_{t-1} = \frac{(1-\alpha_w)(1-\beta\alpha_w)}{\alpha_w(1+\omega_w\theta_w)}\left[\omega_w\widehat{\ell}^h_t - \widehat{\lambda}_t - \widehat{w}_t\right] + \beta\mathrm{E}_t\left\{\widehat{\pi}^w_{t+1} - \gamma_w\widehat{\pi}_t\right\}, \tag{A.4}$$

$$\widehat{\pi}^w_t = \widehat{w}_t - \widehat{w}_{t-1} + \widehat{\pi}_t, \tag{A.5}$$

$$\widehat{y}_t = (1-\alpha)\widehat{\ell}_t + \alpha\widehat{k}_t, \tag{A.6}$$

$$\widehat{\ell}_t = \Omega\widehat{\ell}^h_t, \tag{A.7}$$

$$\widehat{w}_t = \widehat{s}_t + \widehat{y}_t - \widehat{\ell}^h_t, \tag{A.8}$$

$$\widehat{w}^e_t = \widehat{s}_t + \widehat{y}_t, \tag{A.9}$$

$$\widehat{z}_t = \widehat{s}_t + \widehat{y}_t - \widehat{k}_t, \tag{A.10}$$

$$\widehat{\pi}_t - \gamma_p\widehat{\pi}_{t-1} = \frac{\left(1-\alpha_p\right)\left(1-\beta\alpha_p\right)}{\alpha_p}\widehat{s}_t + \beta\mathrm{E}_t\left\{\widehat{\pi}_{t+1} - \gamma_p\widehat{\pi}_t\right\}, \tag{A.11}$$

$$\widehat{x}_t = \widehat{q}_t + \widehat{k}_t - \widehat{n}_t, \tag{A.12}$$

$$\widehat{\check{r}}_t = \mathrm{E}_t\{ R^k_{t+1} \} - \varepsilon_t - \widehat{R}_t, \tag{A.13}$$

$$(x-1)\widehat{b}_t = x(\widehat{q}_t + \widehat{k}_t) - \widehat{n}_t, \tag{A.14}$$

$$\widehat{\check{r}}_{t-1} = f_0 f_1 \widehat{\bar{\omega}}_t + f_3 \sigma_{\omega,t} + \widehat{\pi}_t, \tag{A.15}$$

$$\widehat{\check{r}}_t = \chi \widehat{x}_t + f_5 \mathrm{E}_t \left\{ \sigma_{\omega,t+1} \right\} + \mathrm{E}_t \left\{ \widehat{\pi}_{t+1} \right\} + \varepsilon_t, \tag{A.16}$$

where f_0, f_1, f_3, f_5, and χ are reduced-form parameters;

$$\widehat{n}_t = n_0(\widehat{v}_t + \gamma_t) + \left[1 - n_0\right] \widehat{w}^e_t, \tag{A.17}$$

$$\widehat{v}_t + \widehat{\pi}_t = v_0 \left[\widehat{R}^k_t + \widehat{q}_{t-1} + \widehat{k}_{t-1}\right] - \left[v_0 - 1\right] \left[\widehat{R}_{t-1} + \widehat{b}_{t-1} + \gamma_b \widehat{\pi}_t\right] - v_1 \widehat{\bar{\omega}}_t - v_2 \sigma_{\omega,t}, \tag{A.18}$$

$$\frac{c^e}{k}\widehat{c}^e_t = \frac{c^e}{k}\widehat{v}_t - v_3 \gamma_t, \tag{A.19}$$

where n_0, v_0, v_1, v_2, and v_3 are reduced-form parameters;

$$\widehat{R}^k_t = \widehat{\pi}_t + \widehat{z}_t \left[\frac{z}{r^k}\right] + \widehat{q}_t \left[\frac{1-\delta}{r^k}\right] - \widehat{q}_{t-1} + \left(\widehat{\imath}_t - \widehat{k}_{t-1}\right) \left[\frac{\vartheta \delta^2}{r^k}\right], \tag{A.20}$$

$$\widehat{k}_t = \widehat{k}_{t-1}(1-\delta) + \delta \widehat{\imath}_t, \tag{A.21}$$

$$\widehat{q}_t = \vartheta \delta \left[\widehat{\imath}_t - \widehat{k}_{t-1}\right], \tag{A.22}$$

$$\begin{aligned} \widehat{y}_t = {} & \widehat{c}_t \frac{c}{y} + \widehat{\imath}_t \frac{i}{y} + \widehat{g}_t \frac{g}{y} + \widehat{c}^e_t \frac{c^e}{y} + \left[\widehat{r}^k_t + \widehat{q}_{t-1} + \widehat{k}_{t-1}\right] \left[\mu G(\bar{\omega}) r^k \frac{k}{y}\right] \\ & + \widehat{\bar{\omega}}_t \bar{\omega} \mu G_\omega(\bar{\omega}) r^k \frac{k}{y} + \sigma_{\omega,t} \sigma_\omega \mu G_\omega(\bar{\omega}) r^k \frac{k}{y}, \end{aligned} \tag{A.23}$$

$$\widehat{R}^{not}_t = \rho_R \widehat{R}^{not}_{t-1} + (1-\rho_R)\left[a_\pi \widehat{\pi}_t + a_{\Delta_y} \Delta \widehat{y}_t\right]. \tag{A.24}$$

REFERENCES

Amano R., Shukayev, M. (2012). Risk Premium Shocks and the Zero Bound on Nominal Interest Rates. *Journal of Money, Credit and Banking* 44, 1475–1505.

Bernanke, B., Gertler, M., Gilchrist, S. (1999). The Financial Accelerator in a Quantitative Business Cycle Framework. In Taylor, J.B., Woodford, M., eds., *Handbook of Macroeconomics*, 1341–1393. Amsterdam: Elsevier North-Holland.

Bodenstein, M., Erceg, C. J., Guerrieri, L. (2009). The Effects of Foreign Shocks when Interest Rates are at Zero. International Finance Discussion Paper of the Federal Reserve, 983.

Calvo, G. A. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics* 12, 383–398.

Carlstrom, C. T., Fuerst, T. S., Paustian, M. (2012). How Inflationary is an Extended Period of Low Interest Rate. Federal Reserve Bank of Cleveland, Working Paper 12-02.

Carlstrom, C. T., Fuerst, T. S. (1997). Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. *American Economic Review* 87, 893–910.

Carrillo, J. A., Poilly, C. (2013). How Do Financial Frictions Affect the Spending Multiplier During a Liquidity Trap? *Review of Economic Dynamics* 16, 296–311.

Christiano, L. J., Ikeda, D. (2013). Leverage Restrictions in a Business Cycle Model. NBER Working Paper No. 18688. National Bureau of Economic Research.

Christiano, L. J., Eichenbaum, M., Rebelo, S. (2011). When is the Government Spending Multiplier Large? *Journal of Political Economy* 119, 78–121.

Christiano, L. J., Motto, R., Rostagno, M. (2014). Risk Shocks. *American Economic Review* 104, 27–65.

De Fiore, F., Tristani, O. (2012). (Un)Conventional Policy and the Zero Lower Bound. Manuscript.

Del Negro, M., Giannoni, M., Patterson, C. (2012). The Forward Guidance Puzzle. Federal Reserve Bank of New York, Staff Report No. 574.

Dib, A. (2010). Banks, Credit Market Frictions, and Business Cycles. Working Paper 2010-24, Bank of Canada.

Eggertsson, G., Woodford, M. (2003). The Zero Bound in Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity* 34, 139–235.

Erceg, C. J., Henderson, D. W., Levin, A. T. (2000). Optimal Monetary Policy with Staggered Wage and Price Contracts. *Journal of Monetary Economics* 46, 281–313.

Fernández-Villaverde, J. (2010). Fiscal Policy in a Model with Financial Frictions. *American Economic Review*, Papers and Proceedings, 100, 35–40.

Fernández-Villaverde, J., Gordon, G., Guerrón-Quintana, P. A., Rubio-Ramírez, J. (2012). Nonlinear Adventures at the Zero Lower Bound. NBER Working Papers No. 18058. National Bureau of Economic Research.

Fisher, I. (1933). The Debt-Deflation Theory of Great Depressions. *Econometrica* 1, 337–357.

Gerali, A., Neri, S., Sessa, L., Signoretti, F. M. (2010). Credit and Banking in a DSGE Model of the Euro Area. *Journal of Money, Credit and Banking* 42, 107–141.

Gertler, M., Karadi, P. (2011). A Model of Unconventional Monetary Policy. *Journal of Monetary Economics* 58, 17–34.

Gertler, M., Kiyotaki, N. (2010). Financial Intermediation and Credit Policy in Business Cycle Analysis. In Friedman, B. M., Woodford, M., eds., *Handbook of Monetary Economics*, 547–599. Amsterdam: Elsevier North-Holland.

Gilchrist, S., Yankov, V., Zakrajšek, E. (2009). Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets. *Journal of Monetary Economics* 56, 471–493.

Gilchrist, S., Zakrajšek, E. (2011). Monetary Policy and Credit Supply Shocks. *IMF Economic Review*, 59, 195-232.

Jermann, U., Quadrini, V. (2012). Macroeconomic Effects of Financial Shocks. *American Economic Review* 102, 238–271.

Kiyotaki, N., Moore, J. (1997). Credit Cycles. *Journal of Political Economy* 105, 211–248.

Levin, A., López-Salido, D., Nelson, E., Yun, T. (2010). Limitations on the Effectiveness of Forward Guidance at the Zero Lower Bound. *International Journal of Central Banking* 6, 143–189.

Mertens, K., Ravn, M. O. (2011). Leverage and the Financial Accelerator in a Liquidity Trap. *American Economic Review*, Papers and Proceedings, 101, 413–416.

Nakov, A. (2008). Optimal and Simple Monetary Policy Rules with Zero Floor on the Nominal Interest Rate. *International Journal of Central Banking* 4, 73–127.

Nolan, C., Thoenissen, C. (2009). Financial Shocks and the US Business Cycle. *Journal of Monetary Economics* 56, 596–604.

Oda, N., Takashi, N. (2008). On the Function of the Zero Interest Rate Commitment: Monetary Policy Rules in the Presence of the Zero Lower Bound on Interest Rates. *Journal of the Japanese and International Economies* 22, 34–67.

Smets, F., Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review* 97, 586–606.

Townsend, R. M. (1979). Optimal Contracts and Competitive Markets with Costly State Verification. *Journal of Economic Theory* 21, 265–293.

VI. TAXES, TRANSFERS AND THE MACROECONOMY

JORGE ALONSO-ORTIZ
Instituto Tecnológico Autónomo de México

VI.1 INTRODUCTION: THE MACROECONOMIC IMPORTANCE OF TAXES AND TRANSFERS

Since WWII, the size of government has been increasing around the world. Not only has it increased due to the expenses of the countries involved in the great war (and later efforts on reconstruction), but it also due to the creation and expansion of the welfare state. Although the foundations of the welfare state go back to the late 19th century, dramatic increases in its share of GDP is a more recent phenomenon. These dramatic increases have come hand in hand with increases in taxation, in particular income taxes, payroll taxes, and consumption taxes.

The first social security law was signed in Germany in 1883, which established sickness insurance for workers. France established a voluntary unemployment insurance scheme in 1905. Around the same time, the United States introduced work injury coverage. However, the Social Security Act was not signed until 1935, which established the first U.S. social security system. As time went by, many countries followed and introduced generous and comprehensive insurance schemes.

Today, public social expenditure accounts for more than 20% of GDP across OECD countries. It also accounts for roughly half of total government expenditure. In Section VI.1.1 below, I present a number of stylized facts that will allow the reader to grasp the important role of government taxes and transfers at the macroeconomic level. While I only present aggregated data in this chapter, one should bear in mind that behind this aggregation lies a huge degree of heterogeneity in the institutions that give rise to these aggregate numbers. As we know, "the devil is in the detail".

From all the different types of government expenditures, I will focus solely on social expenditure, as it is the most likely to affect people's labor market behavior directly, and most research analyzes this expenditure item.[1,2] One item that will

[1] I will also present some evidence on the importance of public employment as a transfer program.

[2] By social expenditure I mean social security, unemployment, health, and cash transfers in general.

be ignored is spending on education. While public education is a very important component of government expenditure, and it is important in understanding cross-country macroeconomic performance, its inclusion would make the discussion too sparse and diverse. Consequently, this could make us miss the main point of this chapter, which is that taxes and transfers exert a powerful source of distortion over competitive labor markets, which may, or may not, increase the welfare of people in different countries.

VI.1.1 Stylized facts on taxes, transfers and labor supply

For OECD member countries, the following stylized facts on taxes, transfers, and labor supply have been identified.

1. *Average tax rates have been increasing over time*

   McDaniel (2007) provides a methodology to compute average tax rates across OECD countries. Using her numbers we see that taxes on income (including payroll taxes) in core European countries, Scandinavian countries, the U.K., and the U.S. have risen from 13% to 37%, 15% to 34%, 16% to 27%, and 10% to 14%, respectively, between the period 1950 and 2010. Similarly, taxes on consumption have risen during the same period for the same countries and in the same order from 12% to 19%, 13% to 29%, 12% to 18%, whereas in the U.S. they have remained flat. Corporate taxes have been flat or decreasing over time for all countries in the sample.

2. *There are large differences in average tax rates across countries, regardless of the period of time considered*

   For example, in 2005 the average tax over income and wages was 41% in Belgium, whereas it was only 20% in the U.S.. Similarly, the average tax rate over consumption in 2005 was 31% in Sweden, while in the U.S. it was only 7.5%. Considering that public debt has also been rising for these countries, the share of tax and transfer programs have been increasing dramatically over time. A lower bound can be set if we assume that transfers have to adjust to taxes every period. The share of tax revenues over GDP averaged 35% across OECD countries, with a minimum of 17% for Mexico and a maximum of 48% for Denmark. The actual figure would be even bigger if the stock of public debt had been properly accounted for. In particular, if financial obligations to future cohorts of the population entitled to social transfers were taken into account.

3. *There are large differences across countries in the progressivity of income tax schedules*

The differences discussed above in points 1. and 2. are even more striking when we take also into account the progressivity of transfers (such as social security and unemployment insurance). In countries like Canada, families earning less than 10% of the average income of the economy, receive almost 20% of extra income through subsidies. A family earning more than twice the average income in the economy has to pay roughly 26% of income in taxes, according to OECD Taxing Wages data. This is the rule rather than the exception across OECD countries.

On top of progressivity of taxes we also have to add progressivity of transfers, since omitting this feature may misguide our analysis on the effects of taxes and transfers. Figure VI.1 illustrates a striking example using U.K. and U.S. data. Both countries have similar tax schedules, which would lead us to conclude that the shape of taxes cannot account for differences in labor market outcomes. However, when the progressivity of transfers is taken into account, we get a very different picture. The U.K. subsidizes the income poor significantly more than the U.S. does.

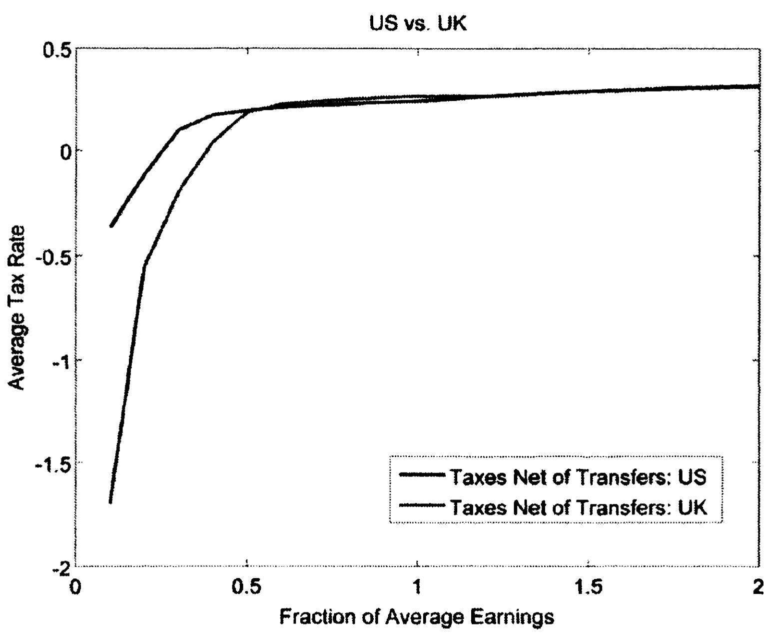


Figure VI.1. Tax and transfer schedules: U.S. vs. U.K.

4. *The majority of tax revenues are raised using distortionary taxation levied on the household*

   Taxes on income and wages account for two thirds of government revenue and this number has remained fairly constant since the 1980's.

5. *The share of social expenditure over total government expenditure has been rising over time*

   During the period 1980–2010, the OECD average for the share of social expenditure over total government expenditure has gone up from 42% to 49%. It has also risen from 18% to 24% as a share of GDP on average. Most of this additional expenditure is due to higher social security receipts, with policy changes and demographics being the major culprits. However, there are other social expenditure programs that may also be important for people decisions, such as disability insurance and unemployment insurance.

   There are also large differences in the size of social expenditures across countries, with Mexico at the bottom of the group with social expenditure accounting for 8% of GDP, and France at the top with social expenditure accounting for 32% of GDP. These differences are huge when I convert these numbers into per capita PPP (U.S. dollars, 2000). In 1980, average expenditure per person-year was 3600, whereas this number was 7600 in 2009. When we look at the cross-sectional distribution of per capita expenditure, we see a similar ranking of countries with Mexico (at the bottom) spending 1200 and Norway (at the top) spending 10700, a ten-fold difference.

   Even though more than half of government expenditure is accounted for by social expenditure, there are other important items. I will consider two: public employment and final government consumption. This last item is defined by the OECD national accounts manual as all the expenditure incurred to produce in kind transfers and public goods and services.[3]

6. *Expenditure on wages of public employees accounts for a quarter of total government expenditure*

   This item accounts for half of all social expenditure over GDP on average, and it equals in importance with expenditure on education or healthcare. While it has remained constant over time, there are significant cross-country differences. For example, Japan spends around 15% of total government expenditure on public salaries, whereas Sweden spends approximately 30%.

[3] The use of the term *wasteful* is frequently used, but it does not imply that people in a country do not value it. It only indicates that it is a subtraction from the expenditure side of the national accounts.

7. *Governments are, on average, the biggest employers*

   Civil service contracts usually have a low lay-off probability and a rigid employment ladder, so they may be an important source of distortions in the labor market. Of course, there are large differences in the share of public employment over total employment across OECD countries. For example, the Japanese civil service accounts for 7% of the employed population, whereas Scandinavian countries average 29%.

8. *Final government expenditure has been increasing over time*

   Final government expenditure has risen on average from 19% to 22% during the period 1980–2010. There are large differences across OECD countries as well, with Mexico (at the bottom) spending 12% of GDP in 2010 and Denmark (at the top) spending 30% of GDP.

   Given the large size of taxes and transfers and its cross-country differences, a natural assumption to make is that there exists a structural relation between tax and transfer programs and hours of work in the aggregate, over the life-cycle, and across countries.

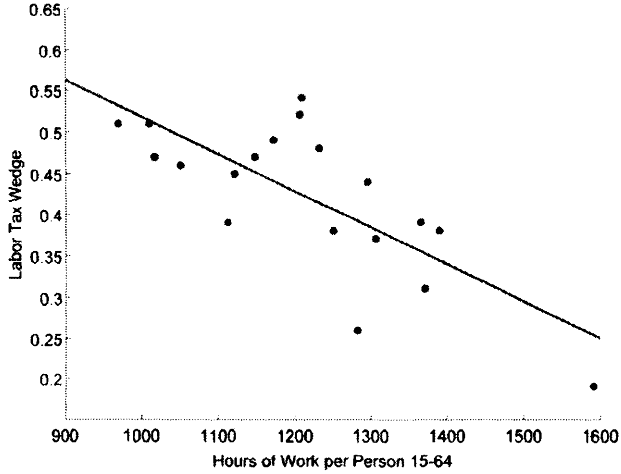
9. *There are large differences across countries, and along time, in hours of work*

   This is well documented in the literature and I refer the reader to the existing literature. See, e.g., McGrattan and Rogerson (2004), Rogerson (2006), Ohanian et al. (2008), Ohanian and Raffo (2012).

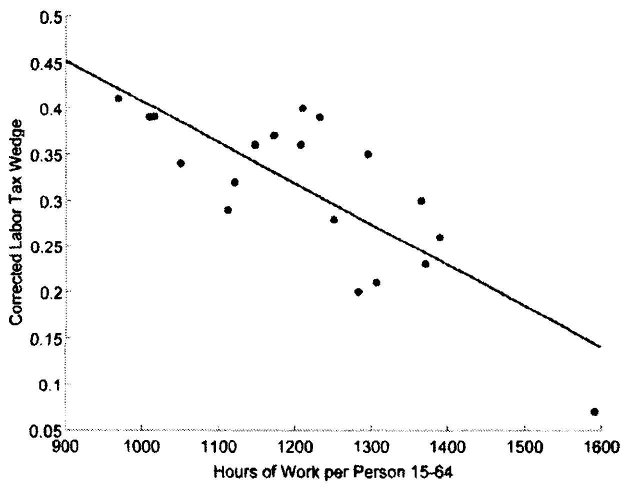
10. *There is a negative relation between hours of work per person (aged 15-64) and taxes*

    This correlation holds independently of the position of each country over the business cycle and is illustrated in Figure VI.2. This is relevant given the recent financial crisis. Using data for available countries in 2009, the correlation of this measure of hours of work and a measure of tax distortions on labor markets known as the *tax wedge* is −0.35. When I correct for final government expenditure this correlation drops to −0.41, as theory predicts. Note that these are unconditional correlations. It is also very important to note that I am using total hours of work, despite a huge literature that treats hours of work of men separately from hours of work of females.[4]

[4] The literature shows that men and women have very different labor market behavior along time and across countries.



(a) Taxes



(b) Taxes less Govt. Expenditure

Figure VI.2. Tax distortions and hours of work per person aged 15–64

11. *There are large cross-country differences in labor supply over the life-cycle*

These differences are particularly big for people approaching retirement age across countries, although cross-country differences in hours of work of young people are also large.

12. *Cross-country differences of hours of work are highly correlated with cross-country differences in social security features, such as entitlement age, replacement rate, and mean tests over social security earnings*[5]

    Both facts 11. and 12. would take a lot of space to document in detail. Interested readers should consult Alonso-Ortiz (2010) and Erosa et al. (2012) who describe the main features of social security programs and how they relate to differences in labor supply at an older age. Next, I will review some papers of this expanding literature.

VI.1.2 Literature summary

Since the beginning of 2000's, a growing literature building on Prescott (2004) has been trying to understand the role that taxes and transfers play on the allocation of aggregate resources. Prescott (2004) starts from the observation that during the 1970's, hours of work per person was relatively high in countries like France, Germany, Japan, and the U.K. , compared to the U.S.; whereas hours of work per person was relatively low in these countries during the 1990's. Even though there could be many different explanations for this phenomena, during the 1970's taxes used to be lower in European countries relative to the U.S., whereas the opposite was true for the 1990's. Using the neoclassical growth model Prescott (2004) finds that taxes can be important in explaining cross-country differences in hours of work.

There have been two main critics to Prescott's work. One strand of the literature argues that in order to get big effects for cross-country differences in hours of work, an implausibly high intertemporal elasticity of labor supply (IES) is needed. Influential papers using aggregate data (see, e.g., Hansen and Singleton, 1983; Hall, 1988) find that intertemporal labor decisions are insensitive to changes in interest rates, implying that the IES is close to zero. Other papers from the microeconometric literature find a low IES that would not back up Prescott's results (see, e.g., MaCurdy, 1981; Browning et al., 1985; Altonji, 1986). To rebut this criticism, it has been argued that the micro IES need not be the same as the macro IES (see, e.g., Imai and Keane, 2004; Prescott, 2004; Guvenen, 2006; Rogerson and Wallenius, 2009; Erosa et al., 2011; Keane and Rogerson, 2011). The core of this rebuttal consists off separating total labor hours into an *intensive margin* (hours per worker) and an *extensive margin* (employment rate). These studies argue that microeconomic research has primarily focused on the behavior of prime

[5] Mean tests are taxes over social security earnings that accrue if a person wants to work and collect social security benefits at the same time.

age males, where there is not much variation in labor supply. However, when labor force participation decisions are considered (such as retirement decisions) an IES consistent with the the micro literature, implies a large IES at the macro level, restating Prescott's results.

The second critique comes from observing Scandinavian countries: if high taxes are related to low hours of work per person, why do countries like Sweden, Norway, and Denmark, who have the highest tax rates among OECD countries, also have hours of work in the upper range relative to the U.S.? In Prescott (2004) taxes are rebated lump sum but in the real world they are spent in a very different way. There is growing evidence that how you spend taxes is key to understand how many hours of work per person a country supplies (Rogerson, 2007). As it will come clear in the next section, whether taxes are spent on social security, production of public goods, or hiring public employees may result in a very different allocation of labor compared to lump-sum rebates.

A different concern that may arise is that the neoclassical growth model may not be the most appropriate modeling framework to quantify the role of taxes in order to explain cross-country differences in hours of work. A relevant departure from the neoclassical growth model takes into account that people suffer uninsurable idiosyncratic risk during their working lives (Aiyagari, 1994). With precautionary savings, people have to decide to what extent they use asset markets or labor supply to insure against this risk (see, e.g., Flodén and Lindé, 2001; Domeij and Flodén, 2006; Pijoan-Mas, 2006; Low et al., 2010). The use of either instrument depends on the IES and the amount of uninsurable idiosyncratic risk people face. The higher the IES, it is expected that people will rely more on changing labor market behavior to cope with idiosyncratic risk.

An important caveat is whether the risk that seems to be observed by the econometrician is the true risk to the individual. There is a long literature trying to back up how much labor income risk is unpredictable.[6] A general finding is that the larger the amount of uninsurable risk, the larger the scope for precautionary savings or working longer hours to cope with it.

The structure of the rest of the chapter is as follows. Section VI.2 outlines a basic framework to help understand how taxes, transfers, and labor supply relate in general equilibrium. Section VI.3 first discusses the role of idiosyncratic labor income risk in an infinite horizon setting to evaluate the impact of taxes and transfers on labor supply, and then presents a life-cycle model with idiosyncratic labor income risk in order to understand how different transfer programs impact on labor supply in the intensive and the extensive margin. Finally, Section VI.4 concludes.

[6] For an in-depth literature review see, e.g., Guvenen (2009), Guvenen and Smith (2010).

VI.2 Taxes, transfers and equilibrium hours of work

The aim of this section is to show that taxes can have a big effect on labor supply. Furthermore, how these taxes are spent will prove key to understanding cross-country differences in labor supply (as in Rogerson, 2007). I begin by illustrating these relationships using a very simple model.

VI.2.1 A static model of general equilibrium

Suppose an economy is populated by a representative firm, a government, and a representative household. The representative firm is characterized by a technology that uses labor as its only input. For expositional purposes, I will assume a Cobb-Douglas function

$$y = Al^{1-\alpha}. \qquad \text{(VI.1)}$$

Given this technology, the objective of the firm is to maximize profits (π^*), where all firms are owned by the household.

In the real world there are many different taxes and transfer programs, which follow complicated rules that cannot be modeled using this simple framework. However, despite its simplicity, it will be shown that this model is able to accurately predict labor supply behavior across countries.

The government sets taxes and transfers. The government can use lump-sum taxes (T), income taxes (τ_y), and consumption taxes (τ_c) in any combination. Taxes on income are proportional and taxes on consumption are *ad valorem*. From the transfers side, the government may distribute tax proceeds as lump-sum transfers back to the household (Ω). It may also buy private consumption (G) to produce a public good or it may throw it away as non-productive bureaucratic expenses, inefficiency, or even corruption.[7] In addition, the government may also hire public employees. We assume that if the government demands public employment (L_g) it pays the equilibrium wage rate. Given these tax and transfer programs, the government is defined by the following budget constraint:

$$T + \tau_c c^* + \tau_y y^* = gy^* + w^* L_g + \Omega. \qquad \text{(VI.2)}$$

Note that this budget constraint subsumes all types of expenditures outlined in the previous section. For example, in this model social security transfers would be lump-sump transfers, and government expenditure may include health expenditure, expenditure on materials, investment in public capital, and so on.

[7] To obtain analytical solutions I assume $G = gy$.

The representative household is characterized by a utility function

$$u(c, h, G) = \gamma \log(h) + \log(c) + \Psi \log(G), \qquad \text{(VI.3)}$$

that depends on consumption (c), leisure (h), and government expenditure (G). The household maximizes utility subject to the following budget constraint:

$$(1 + \tau_c)c = (1 - \tau_y)w(l + L_g) + (1 - \tau_y)\pi^* - T + \Omega. \qquad \text{(VI.4)}$$

From solving the representative household problem, we obtain a key relation between the marginal relation of substitution and net wages:

$$\frac{(1 + \tau_c)c^*}{H - l^* - L_g} = (1 - \tau_y)(1 - \alpha)\frac{y^*}{l^*}, \qquad \text{(VI.5)}$$

where $H = h + l^* + L_g$ is the total time available for leisure and work (l^*). I imposed clearing in the labor market, which guarantees clearing in the product market as well. Solving equation (VI.5) for equilibrium labor supply l^* provides a fundamental relation between hours of work, and taxes and transfers:

$$l^* = \frac{(1 - \alpha)(H - L_g)}{1 - \alpha + \gamma \frac{(1+\tau_c)(1-g)}{1-\tau_y}}. \qquad \text{(VI.6)}$$

Note that we can write down all distortions, except public employment, into a sufficient statistic that it is commonly known as the *labor wedge*:

$$1 - \tau = \frac{1 - \tau_y}{(1 + \tau_c)(1 - g)}. \qquad \text{(VI.7)}$$

It is also possible to study the behavior of total hours of work including people working in the public sector:

$$L^* = l^* + L_g = \frac{(1 - \alpha)H + \frac{\gamma}{1-\tau} L_g}{1 - \alpha + \frac{\gamma}{1-\tau}}. \qquad \text{(VI.8)}$$

Equation (VI.7) contains the main testable implications: countries with a larger labor wedge will supply less hours of work than in the U.S., with two caveats. If a given country has a larger share of final government expenditure, or a larger size of public employment relative to the U.S., this (relatively) increases its labor supply. Figure VI.3 shows how well this simple model predicts cross-country labor supply relative to the U.S., against data for 2005. The scatter plot presents predictions for three different computations of the labor wedge. As this is a static model, I first consider that taxes on investment are consumption taxes and taxes

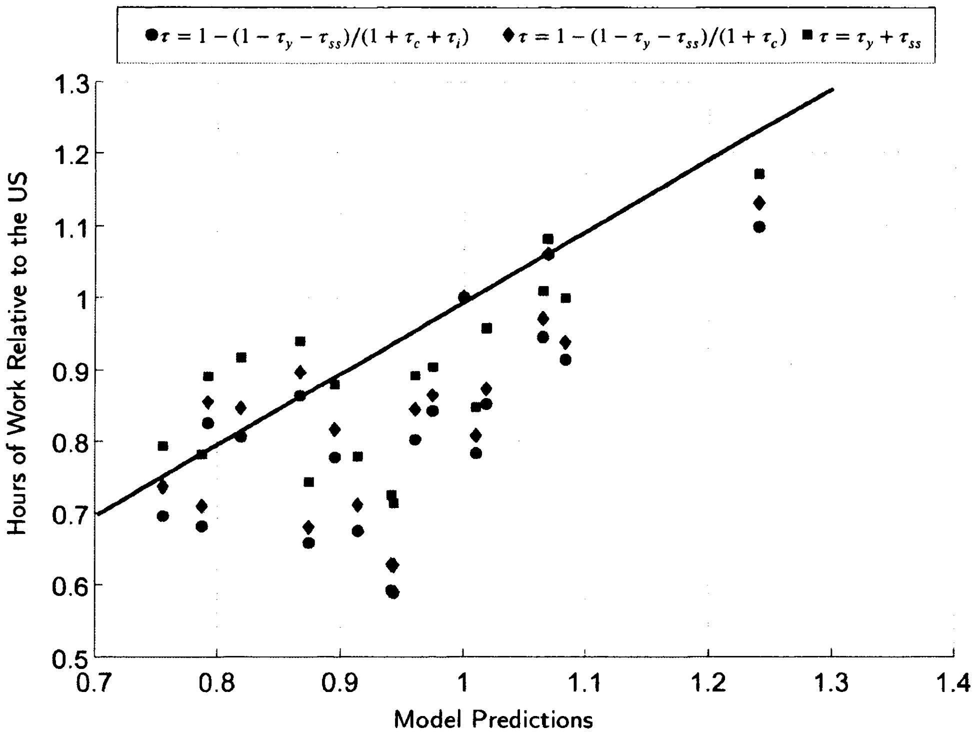


Figure VI.3. Model predictions

on capital income are taxes on profits and labor income. Finally, I take into account taxes that directly distort labor markets. When broad definitions of the labor wedge are considered, the model systematically over-predicts labor supply relative to the U.S.. This arises because in this simple model there is no capital which is also distorted. As capital and labor are complements, abstracting from capital contributes to the over-prediction of labor supply.

VI.2.2 Prescott (2004)

Prescott (2004) derives his results using the neoclassical growth model. Compared to the static model of Section VI.2.1, this model allows us to study economies with distortions on capital, and economies that may not be on the balanced growth path at all times.

Technology is characterized by a Cobb-Douglas production function

$$y_t = k_t^{\alpha}\left(A_t l_t^{1-\alpha}\right),$$

that is used by a representative firm to produce a homogeneous product that can be consumed or invested:

$$y_t = c_t + i_t.$$

Capital accumulates according to a geometric law of motion

$$k_{t+1} = (1 - \delta)k_t + i_t,$$

where $0 < \delta < 1$ is the depreciation rate. In this economy there is a representative consumer that maximizes an infinite flow of utility

$$u(c, h, G) = \sum_{t=0}^{\infty} \left\{\gamma \log(h_t) + \log(c_t) + \Psi \log(G_t)\right\},$$

where $H_t = h_t + l_t + L_t^g$, subject to a budget constraint

$$(1 + \tau_t^c)c_t + (1 + \tau_t^i)i_t = (1 - \tau_t^y)(w_t(l_t + L_t^g) + r_t k_t) + \Omega.$$

The equilibrium in this economy is unique and determined by an intertemporal efficiency condition (equivalent to equation (VI.5)), an Euler equation, the government budget constraint, and the resource constraint. Prescott (2004) uses equation (VI.5) only to map the economy's data into a simple equation which holds out of the balanced growth path of any economy. In his work, he abstracts from public employment and government expenditure. Equation (VI.9) shows the accounting relation that predicts cross country differences in hours of work, given differences in the tax wedge:

$$\frac{l_{it}}{H_{it}} = \frac{1 - \alpha}{1 - \alpha + \frac{c_{it}}{y_{it}} \frac{\gamma}{1-\tau_{it}}}. \tag{VI.9}$$

Assuming preferences are identical across countries, Prescott (2004) shows that a careful measure of tax wedges and the consumption-output ratio would predict differences in hours of work across countries. Based on the work of Mendoza et al. (1994), he constructs a measure of tax wedges that is consistent with his model.[8] Table VI.1 summarizes his findings.

It is striking how well differences in tax wedges, when tax proceeds are rebated lump-sum, are able to predict differences in hours of work, conditional on the consumption-output ratio. The intuition is that in a perfect foresight model the consumption-output ratio is a sufficient statistic of the distribution of future states. This accurateness is something that would not have achieved if countries were assumed to be on the balanced growth path.

[8] McDaniel (2007) provides improved measures of tax rates that can be used to calculate labor wedges. The results hold under both measures.

Table VI.1. Actual and predicted labor supply

		Labor supply		
Period	*Country*	*Actual*	*Predicted*	*Error*
1993–1996	Germany	19.3	19.5	−0.2
	France	17.5	19.5	−2.0
	Italy	16.5	18.8	−2.3
	Canada	22.9	21.3	1.6
	UK	22.8	22.8	0.0
	Japan	27.0	29.0	−2.0
	US	25.9	24.6	1.3
1970–1974	Germany	24.6	24.6	0.0
	France	24.4	25.4	−1.0
	Italy	19.2	28.3	−9.1
	Canada	22.2	25.6	−3.4
	UK	25.9	24.0	1.9
	Japan	29.8	35.8	−6.0
	US	23.5	26.4	−2.9

Notes: Labor supply is measured as weekly hours worked per person aged 15 – 64.
Source: Prescott (2004).

The Euler equation is not used at all in Prescott (2004) but it can be used to do development accounting. It is a well documented fact (see, e.g., Hall and Jones, 1999; Caselli, 2005; Hsieh and Klenow, 2010) that differences in labor and capital are able to account for differences in the GDP per capita of rich countries. However, these differences cannot account for differences in the world income distribution as their order of magnitude is 40 (Parente and Prescott, 1993). Restuccia and Urrutia (2001) use the neoclassical growth model to study how differences in taxes account for differences in savings rates. They show that these rates can be predicted with precision, even though taxes cannot account for differences in GDP per capita across countries.

What we learn from this exercise is that differences in taxes and transfers account for differences in the allocation of aggregate factors of production: labor and capital. Will these results change if idiosyncratic labor income risk and incomplete markets are introduced? This is the focus of the next section.

VI.3 TAXES AND TRANSFERS UNDER UNCERTAINTY AND INCOMPLETE MARKETS

The neoclassical growth model assumes a representative agent but in the real world we observe a lot of heterogeneity along many dimensions (e.g., income and wealth). This heterogeneity may have implications for the labor supply and savings decisions taken by risk averse people, particularly if there are no markets to trade contingent claims in every state of nature. Huggett (1993) acknowledges that this feature may be key to account for the equity premium puzzle (Mehra and Prescott, 1985). Similarly, Aiyagari (1994) writes down a model where individuals suffer from uninsurable idiosyncratic risk, they are subject to borrowing constraints, and markets are incomplete. He realizes that these features may account for why people "over-save" with respect to what the neoclassical model would predict. Pijoan-Mas (2006) extends the model of Aiyagari (1994) to study the implications for labor supply decisions. The importance of this extension is that people may not only use savings decisions to self-insure against risk, but they may choose to work longer hours when they get an array of good shocks to insure against the possibility of an array of bad shocks.[9] They may also choose to work longer hours in the face of an array of bad shocks to compensate for lost labor income, as they would be able to supply less efficiency units of labor for the same number of hours worked. The quantitative importance of these mechanisms depends crucially on two things: the IES and the persistence of shocks.

By interacting in a non-trivial way with risk and people's labor supply and savings decisions, the role of taxes and transfers in this set up may change. For example, differences in the progressivity of taxes and transfers change the amount of risk that someone is experiencing ex-post, thereby affecting their choices on how much to save and work. Also, the manifold of differences in the rules that govern the allocation of transfers may be key to account for cross-country differences in labor supply and differences in its distribution over the life-cycle. There are many transfer programs that kick in only in cases where income falls below a consumption floor, or in social security programs in which pension accrual depends on many individual features, such as age, income, assets, marital status, labor marked duration spells, and so on.

[9] These array of bad shocks will happen with probability one in an infinite horizon model as the idiosyncratic risk stochastic process shows mean reversion.

VI.3.1 Precautionary savings or working longer hours

Consider an economy populated by a continuum of ex-ante identical individuals of measure one. Individuals live forever, and when the are born into the economy they suffer from a realization of an idiosyncratic shock to their efficiency units of labor,[10] z, where

$$\log(z') = \rho \log(z) + \epsilon', \tag{VI.10}$$

$$\epsilon' \overset{i.i.d.}{\longrightarrow} N(0, \sigma_\epsilon^2),$$

and the prime symbol refers to variables next period. Individuals make labor supply (l), consumption (c), and asset accumulation decisions (a'), given their state variables: assets (a) and a idiosyncratic shock to labor productivity (z). The recursive representation of the individual's decision problem is:

$$v(a, z) = \max_{c,l,a'} \left\{ \frac{c^{1-\sigma} - 1}{1 - \sigma} + \lambda \frac{(1 - l)^{1-\nu} - 1}{1 - \nu} + \beta \mathrm{E}_{z'|z} \left[v(a', z') \right] \right\}, \tag{VI.11}$$

subject to:

$$c + a' = wzl + (1 + r)a - T(c, wzl, ra) + \Omega(c, wzl, ra),$$
$$a' \geq 0, c \geq 0,$$

and subject to (VI.10), where T and Ω are generic tax and transfer programs that depend on consumption, labor income, and savings' returns. These taxes and transfers may be progressive, or they may be proportional taxes coupled with lump-sum transfers as well. Furthermore, these functions are general enough to capture any tax and transfer schedule found across countries. Production takes place via a representative firm that maximizes profits, subject to a standard Cobb-Douglas production function. In equilibrium, product markets, capital markets, and labor markets clear:[11]

$$\int_{A \times Z} a'(a, z) \mu(da, dz) = K, \tag{VI.12}$$

$$\int_{A \times Z} zl(a, z) \mu(da, dz) = L, \tag{VI.13}$$

[10] This shock is uncorrelated across individuals. Thus, there are no insurance opportunities.

[11] After invoking Walras law, we only need to present equilibrium conditions for capital and labor markets.

where $a'(\cdot)$ and $l(\cdot)$ are policy functions for asset accumulation and labor decisions, $\mu(\cdot)$ is a measure of agents with assets a, and a realization of idiosyncratic shock z in the product space of assets and productivities $A \times Z$.[12] Pijoan-Mas (2006) calibrates this model for the U.S. economy. He finds that while the model accounts very well for the distribution of labor income and hours of work across asset quintiles, it is not able to account for the wealth distribution of the U.S.[13]

In addition, Pijoan-Mas (2006) uses this model to quantify the importance of precautionary savings versus working longer hours. In his benchmark calibration, individuals work on average 15% more and hold a stock of capital 18% higher compared to the complete markets version of the model. Therefore, individuals work an inefficiently high number of hours. Key to his results are the IES (ν), and the variance and persistence of idiosyncratic shocks.

VI.3.2 Taxes and transfers in an incomplete markets model

If the amount of risk an individual faces is related to her labor supply decisions, it may also interact with taxes and transfers. Consider, as an illustration, a very simple tax and transfer program in which there is a proportional tax on labor income and tax proceeds are divided equally across people in the economy, regardless of their idiosyncratic characteristics. In such a system, progressivity arises as those who suffered from an array of good shocks to their efficiency units of labor would pay more taxes than what they receive in terms of transfers. Similarly, those under an array of bad shocks would be receiving more transfers than what they have paid in terms of taxes. This feature distorts the amount of risk an individual perceives and it would be expected that their labor supply decisions change.

Alonso-Ortiz and Rogerson (2010) build on this intuition, and using the model laid down in the previous subsection, they quantify whether cross-country differences in taxes account for differences in hours of work as in Prescott (2004). They find that taxes and transfers still explain differences in labor supply. The effect is actually stronger because of the working longer hours mechanism. For example, a change in taxes from 0.30 to 0.50 in the representative agent model generates a drop in hours worked of 21%, whereas in the heterogeneous agent model it falls by 27%. A drawback of the representative agent model is that any distortionary tax reduces welfare. This is no longer true in a heterogeneous agent economy. Alonso-Ortiz and Rogerson (2010) find that the U.S. is very close to the

[12] There are many technical details that I am not covering. The interested reader may consult Aiyagari (1994) or Pijoan-Mas (2006). For a recent review of this literature see Guvenen (2011).

[13] This is a common flaw of this literature. See Castañeda et al. (2003).

welfare maximizing tax rate, so that increasing taxes to Continental Europe levels would reduce welfare. This therefore suggests that Americans are "happier". A final implication of the model is that the apparent productivity catch up of European countries to the U.S. may be an illusion. In the presence of high taxes, only the most productive work, as the least productive would be living on lump-sum transfers. This creates a selection effect that shows up as higher measured labor productivity.

VI.3.3 The life-cycle model

Many transfer programs depend on age, so a model where agents live forever might not be the right abstraction to understand their effects. As has already been documented in Section VI.1 above, the labor supply behavior for specific age groups is markedly different. To understand these features we need a model that includes individual ageing. Huggett (1996) presents a variation of the infinitely-lived heterogeneous agent model, extended to account for overlapping generations. As there are many variations of this model, I will write down a version similar to Erosa et al. (2012), as it will allow me to consider a wide range of different tax and transfer programs. The model will be able to capture differences in labor supply along both the intensive and extensive margin.

Consider a measure one of individuals that are born in the economy with a given age, $j = 1$, and they live to a maximum of J years. As they age, they may die with certain probability that depends on age, p_j. In the steady state of this economy, there will be a constant number of people of each age, η_j. In every year of their lives, people may choose consumption (c), hours of work (l), and assets (a') to maximize their lifetime utility. The individual decision problem can be written down recursively:

$$v_j(a,z,h) = \max_{c,l,a'} \left\{ u_j(c,l,h) + \beta p_{j+1} \mathrm{E}_{z'|z,h'|h} \left[v_{j+1}(a',z',h') \right] \right\}, \qquad \text{(VI.14)}$$

subject to:

$$c + a' = w(l)zl + (1+r)a + B - T_j(c, w(l)zl, ra, h) + \Omega_j(c, w(l)zl, ra, h),$$

and $a' \geq 0$. This representation is very similar to equation (VI.11) except that the Bellman equation depends on a larger set of state variables. To make individual decisions interesting enough I extend the previous state space to include age and health (h).[14] Taxes and transfers depend on consumption, labor income, assets

[14] Other relevant states would be: human capital, marital status, number of children, whether the individual has claimed disability insurance or social security, etc...

income, health, and age. These generic functions may capture consumption floor subsidies like food stamps in the U.S., health related tax deductions, disability insurance, and social security benefits (including health benefits.) Another important difference with the previous model is that wages per efficiency unit of labor depend on hours of work. Aaronson and French (2004) document empirically a discontinuity in the return to labor services. They actually find that labor services' pricing function is non-convex. This non-convexity is the mechanism used by Erosa et al. (2012) to generate extensive margin labor market decisions. Finally, B stands for accidental bequests as people in this economy may die leaving unused assets with a positive probability. For simplicity, it is assumed that the government collects these assets and distributes them lump-sum to all agents in the economy.[15]

There is a representative firm that produces an homogenous product using a slightly different than usual Cobb-Douglas function

$$Y = l^{\epsilon} K^{\alpha}(Az)^{1-\alpha},$$

where $0 < \epsilon \leq 1$. If $\epsilon = 1$, workers are not subject to fatigue, otherwise increasing the work-week length will be subject to the law of diminishing returns. With $\epsilon = 1 - \alpha$ technology collapses to the standard Cobb-Douglas, but if $\epsilon > 1 - \alpha$ hours and effective labor are imperfect substitutes and their composition may matter. The representative firm maximizes profits choosing capital only. In this set up, profits may be positive which is at odds with the idea of reproducibility of a plant at the aggregate level. Introducing a free entry condition to firms rules out this possibility. Under free entry, I can derive the following pricing function for hours of work:[16]

$$w(l) = r\frac{1-\alpha}{\alpha}\left[\frac{\alpha l^{\epsilon}}{r}\right]^{\frac{1}{1-\alpha}}.$$

With this pricing function, individuals will have the incentive to withdraw from the labor force if they want to work less than a certain threshold l^*, generating behavior resembling retirement decisions.[17] What is left is to examine the quantitative results of this model in order to evaluate tax and transfer programs.

[15] However, there are many papers that deal with bequest and altruism in detail. See Fuster et al. (2008) as an example and for further references.

[16] See Erosa et al. (2010) for details.

[17] There are other ways to generate retirement decisions, either by focussing on the extensive margin only, as in Alonso-Ortiz and Rogerson (2010) or Alonso-Ortiz (2010), or introducing some utility cost of participating in the labor market, as in French (2005) and French and Jones (2011).

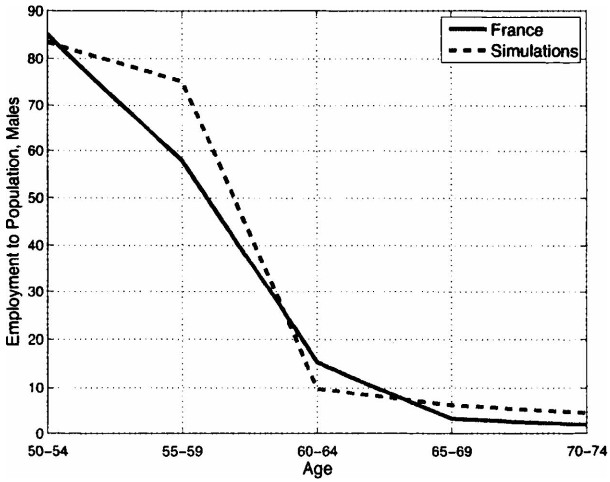
VI.3.4 Effects of tax and transfer programs over the life-cycle and in the aggregate

There are a bulk of papers analyzing the impact of taxes and transfers over the life-cycle in many countries. There are also many papers studying the likely effects, in terms of aggregate allocations and welfare, of reforming tax and transfer programs. In particular, there is a large literature that analyzes the effect of social security with its many components around the world (Gruber and Wise, 2004, 2007, 2010) but not so much work has been done on understanding how measurable differences in tax and transfer programs account for differences in life-cycle labor supply, using a parsimonious model. I will focus on a few papers on this last strand of the literature. Alonso-Ortiz (2010) is one of the first papers that uses a version of the model outlined in Section VI.3.3 to quantify the impact of cross-country differences in social security to account for differences in labor supply behavior of people aged 50+. In his paper, three social security features are selected: early entitlement age to the program, average replacement rate,[18] and whether social security entitlement rules allow a person to collect social security and still work. He finds that when such a model is calibrated for the U.S. economy, differences in these three features only account for two-thirds of the differences in retirement behavior across the OECD.[19] Contrary to a commonly held prior, differences in entitlement age play the least important role, whilst differences in the replacement rate and whether you can collect benefits and work at the same time matter most. As Figure VI.4 illustrates for France and Germany, the model is also able to predict quite accurately the life-cycle labor supply of people aged 50+.

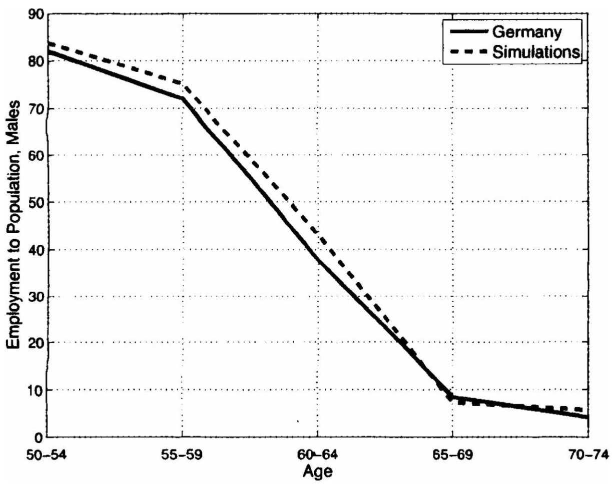
Even though the model works well for females too, there are many other features that affect female labor supply (although it is more likely that these other features kick in early in life and not to females already aged 50+ in the labor force). Erosa et al. (2012) follow a similar path and look at how cross-country differences in social security taxes and disability insurance rules account for differences in labor supply of people aged 50+. Compared to Alonso-Ortiz (2010), they include an intensive margin and a disability risk. Considering disability insurance is not a trivial extension, as people may consider scheduling claims to each program in order to retire earlier while maximizing benefits. Erosa et al. (2012) find that social security taxes explain a substantial amount of cross-country differences in elderly labor supply as in Alonso-Ortiz (2010). Surprisingly as well, they do not

[18] This is defined as the percentage of a measure of previous lifetime average wages that is paid as a pension every period, once benefits are claimed.

[19] This is done to fix all variables to U.S. levels and be able to perform a proper comparative statics analysis.



(a) France



(b) Germany

Figure VI.4. Labor supply of men aged 50+

find big effects of disability insurance except for a few countries. However, this is not the last word on the interaction of social security and disability insurance as single country evidence points in the opposite direction. Kitao (2012) shows that the interaction between unemployment, health, disability insurance, and social security is far from trivial. This goes in the line of previous work that puts together social security and health insurance, which finds that both are key to understand retirement behavior (see, e.g., Rust and Pheland, 1997; De Nardi et al., 2010).

Finally, these type of life-cycle models do not only have implications for labor supply over the life-cycle. A suitable example of this is Guvenen et al. (2014), who use a version of the model presented here, extended to account for human capital accumulation. Guvenen et al. (2014) document an increase in wage inequality in the U.S. compared to continental European countries. Why are there cross-country differences in wage inequality, as well as differences over time? Guvenen et al. (2014) find that differences in the progressivity of the tax code may be behind the empirical evidence they present. A progressive tax code compresses ex-post the wage distribution. If there are not big wage gains at stake from human capital formation then people will have less incentive to accumulate human capital, reducing ex-ante wage inequality. The model also seems to be consistent with cross country differences in hours of work, as in Prescott (2004), but their model enables them to understand the mechanisms that generate such results.

VI.4 Conclusions

This chapter intends to bring interest to the reader about the analysis of taxes and transfers using models of general equilibrium. This is a fast growing area of research that lies in the intersection of macroeconomics, labor economics, and econometrics. Given space constraints, it is not possible to explore the relation between these three areas, but I hope to have succeeded into providing some background so the that reader can explore and, more importantly, formulate relevant research questions.

I hope that it is noticeable from the introduction that there are more stylized facts than theories to account for them. This is because there is still a lot to be understood on the effect that taxes and transfers may have over our life-cycle and in the aggregate.

Finally, I would like to emphasize that while all the evidence presented here pertains to OECD countries, there is also a lot of work to be done in Latin American countries as well; both in terms of documenting the facts, and proposing theories to account for them. These countries present many challenges, due to the presence of a sizable informal sector and widespread corruption. One question that may arise, in light of the topics presented in this chapter, concerns the relationship between taxes, transfers, and the size of the informal sector. I believe this is not only a promising area of research, but a responsibility of Mexican economists, as the use of such models might enlighten policymakers to implement better policies.

REFERENCES

Aaronson, D., French, E. (2004). The Effect of Part-Time Work on Wages: Evidence from the Social Security Rules. *Journal of Labor Economics* 22, 329–352.

Aiyagari, S.R. (1994). Uninsured Idiosyncratic Risk and Aggregate Saving. *Quarterly Journal of Economics* 109, 659–684.

Algan, Y., Cahuc, P., Zylberberg, A. (2002). Public Employment: Does it Increase Unemployment? *Economic Policy* 17, 7–66.

Altonji, J. G. (1986). Intertemporal Substitution in Labor Supply: Evidence from Micro Data. *Journal of Political Economy* 94, 176–215.

Alonso-Ortiz, J. (2010). Social Security and Retirement Across the OECD. Mimeo, ITAM.

Alonso-Ortiz, J., Rogerson, R. (2010). Taxes, Transfers and Employment in an Incomplete Markets Model. *Journal of Monetary Economics* 57, 949–958.

Auerbach, A. J., Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.

Browning, M., Deaton, A., Irish, M. (1985). A Profitable Approach to Labor Supply and Commodity Demands over the Life-Cycle. *Econometrica* 53, 503-543.

Caballé, J., Fuster, L. (2003). Pay-as-you-go Social Security and the Distribution of Altruistic Transfers. *Review of Economic Studies* 70, 541–567.

Caselli, F. (2005). Accounting for Cross-Country Income Differences. In Aghion, P., Durlauf, S., eds., *Handbook of Economic Growth*, 679–741. Amsterdam: Elsevier North-Holland.

Castañeda, A., Diaz-Gimenez, J., Rios-Rull, J. (2003). Accounting for the U.S. Earnings and Wealth Inequality. *Journal of Political Economy* 111, 818–857.

Chang, Y., Kim, S. (2006). From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agents Macroeconomy. *International Economic Review* 47, 1–27.

Conesa, J. C., Krueger, D. (2006). On the Optimal Progressivity of the Income Tax Code. *Journal of Monetary Economics* 53, 1425–1450.

De Nardi, M., French, E., Jones, J. B. (2010). Why do the Elderly Save? The Role of Medical Expenses. *Journal of Political Economy* 118, 39–75.

Domeij, D., Flodén, M. (2006). Population Aging And International Capital Flows. *International Economic Review* 47, 1013–1032.

Erosa, A., Fuster, L., Kamburov, G. (2010). The Heterogeneity and Dynamics of Individual Labor Supply over the Life Cycle: Theory and Evidence. Mimeo.

Erosa, A., Fuster, L., Kamburov, G. (2011). Towards a Micro Founded Theory of Aggregate Labor Supply. IMDEA Working Papers Series in Economic and Social Sciences 2011/13.

Erosa, A., Fuster, L. Kamburov, G. (2012). Labor Supply and Government Programs: A Cross-Country Analysis. *Journal of Monetary Economics* 59, 84–107.

Flodén, M., Lindé, J. (2001). Idiosyncratic Risk in the United States and Sweden: Is there a Role for Government Insurance? *Review of Economic Dynamics* 4, 406–437.

Flodén, M. (2005). Labor Supply and Saving under Uncertainty. *The Economic Journal* 116, 721–737.

French, E. (2005). The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *Review of Economic Studies* 72, 395–427.

French, E., Jones, J. B. (2011). The Effects of Health Insurance and Self-insurance on Retirement Behavior. *Econometrica* 79, 693–732.

Fuster, L., Imrohoroglu, A., Imrohoroglu, S. (2008). Altruism, Incomplete Markets, and Tax Reform. *Journal of Monetary Economics* 55, 65–90.

Gruber, J., Wise, D. A. (2004). *Social Security Programs and Retirement around the World: Micro-Estimation*. Chicago: The University of Chicago Press.

Gruber, J., Wise, D. A. (2007). *The Fiscal Implications of Social Security Reform Around the World*. Chicago: The University of Chicago Press.

Gruber, J., Wise, D. A. (2010). *Social Security Programs and Retirement around the World: The Relationship to Youth Unemployment*. Chicago: The University of Chicago Press.

Guvenen, F. (2006). Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective. *Journal of Monetary Economics* 53, 1451–1472.

Guvenen, F. (2009). An Empirical Investigation of Labor Income Processes. *Review of Economic Dynamics* 12, 58–79.

Guvenen, F. (2011). Macroeconomics with Heterogeneity: A Practical Guide. NBER Working Papers No. 17622. National Bureau of Economic Research.

Guvenen, F., Kuruscu, B., Ozcan, S. (2014). Taxation of Human Capital and Wage Inequality: A Cross Country Analysis. Forthcoming, *Review of Economic Studies*.

Guvenen, F., Smith, A. (2010). Inferring Labor Income Risk From Economic Choices: An Indirect Inference Approach. NBER Working Papers No. 16327. National Bureau of Economic Research.

Hall, R. E. (1988). Intertemporal Substitution in Consumption. *Journal of Political Economy* 96, 339–357.

Hall, R. E., Jones, C. I. (1999). Why do Some Countries Produce so Much More Output per Worker Than Others? *Quarterly Journal of Economics* 114, 83–116.

Hansen, L. P., Singleton, K. J. (1983). Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns. *Journal of Political Economy* 91, 249–265.

Hsieh, C., Klenow, P. J. (2010). Development Accounting. *American Economic Journal: Macroeconomics* 2, 207–223.

Huggett, M. (1993). The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. *Journal of Economic Dynamics and Control* 17, 953–969.

Huggett, M. (1996). Wealth Distribution in Life Cycle Economies. *Journal of Monetary Economics* 38, 469–494.

Imai, S., Keane, M. P. (2004). Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review* 45, 601–641.

Keane, M. P., Rogerson, R. (2011). Reconciling Micro and Macro Labor Supply Elasticities: A Structural Perspective. NBER Working Papers No. 17430. National Bureau of Economic Research.

Kitao, S. (2012). A Life Cycle Model of Unemployment and Disability Insurance. Mimeo, Hunter College.

Low, H., Meghir, C., Pistaferri, L. (2010). Wage Risk and Employment Risk over the Life Cycle. *American Economic Review* 100, 1432–1467.

MaCurdy, T. E. (1981). An Empirical Model of Labor Supply in a Life-Cycle Setting. *Journal of Political Economy* 88, 1059–1085.

McDaniel, C. (2007). Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD 1950-2003. Mimeo, Arizona State University.

McDaniel, C. (2011). Forces Shaping Hours in the OECD, 1960-2004. *American Economic Journal: Macroeconomics* 3, 27–52.

McGrattan, E., Rogerson, R. (2004). Changes in Hours Worked, 1950-2000. *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 14–33.

Mendoza, E., Razin, A., Tesar, L. (1994). Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption. *Journal of Monetary Economics* 34, 297–323.

Mehra, R., Prescott, E. C. (1985). The Equity Premium Puzzle. *Journal of Monetary Economics* 15, 145–161.

Ohanian, L. E., Raffo, A., Rogerson, R. (2008). Long Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004. *Journal of Monetary Economics* 55, 1353–1362.

Ohanian, L. E., Raffo, A. (2012). Aggregate Hours Worked in OECD Countries: New Measurement and Implications for Business Cycles. *Journal of Monetary Economics* 59, 40–56.

Parente, S. L., Prescott, E. C. (1993). Changes in the Wealth of Nations. *Federal Reserve Bank of Minneapolis Quarterly Review* 17, 3–16.

Pijoan-Mas, J. (2006). Precautionary Savings or Working Longer Hours? *Review of Economic Dynamics* 9, 326–352.

Prescott, E. C. (2004). Why do Americans Work so Much More than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review* 28, 2–13.

Restuccia, D., Urrutia, C. (2001). Relative Prices and Investment Rates. *Journal of Monetary Economics* 47, 93–121.

Rogerson, R. (2006). Understanding Differences in Hours Worked. *Review of Economic Dynamics* 9, 365–409.

Rogerson, R. (2007). Taxation and Market Work: Is Scandinavia an Outlier? *Economic Theory* 32, 59–85.

Rogerson, R., Wallenius, J. (2009). Micro and Macro Elasticities in a Life-Cycle Model with Taxes. *Journal of Economic Theory* 144, 2277–2292.

Rust, J., Pheland, C. (1997). How Social Security and Medicare Affect Retirement Behaviour in a World of Incomplete Markets. *Econometrica* 65, 781–831.

Contemporary Topics in Macroeconomics
se terminó de imprimir en agosto de 2014
en los talleres de Master Copy, S.A. de C.V,
Av. Coyoacán 1450, col. Del Valle,
03220 México, D.F.
Portada: Pablo Reyna.
Tipografía y formación: Axel Retif.
Cuidó la edición la Dirección de Publicaciones de
El Colegio de México.

This book is a compilation of six self-contained chapters written by expert contributors covering a number of important issues in modern macroeconomics, all employing the techniques of dynamic general equilibrium. The topics covered in this collection include:

"Reflections on modern macroeconomics",
J.C. Leal Ordóñez and S. McKnight

"Business cycles in developed and developing countries",
A. Antón Sarabia

"Designing monetary policy rules for small open economies",
S. McKnight

"The informal sector in contemporary models of the aggregate economy",
J.C. Leal Ordóñez

"Investigating the zero lower bound on the nominal interest rate under financial instability",
J.A. Carrillo and C. Poilly

"Taxes, transfers and the macroeconomy",
J. Alonso-Ortiz

ISBN: 978-607-462-637

9 786074 62637

EL COLEGIO DE MÉXICO