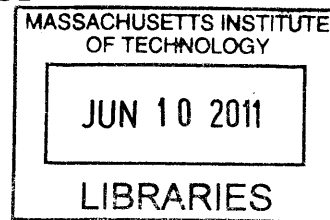


**Financial Distortions and the Distribution of  
Global Volatility**

by  
Maya Rachel Eden  
B.S., Hebrew University (2006)  
Submitted to the Department of Economics  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Economics

at the  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
June 2011

© Maya Rachel Eden, MMXI. All rights reserved.  
The author hereby grants to MIT permission to reproduce and  
distribute publicly paper and electronic copies of this thesis document  
in whole or in part.



**ARCHIVES**

Author .....

Department of Economics

May 15, 2011

Certified by .....

..... Saballero

Ford International Professor of Economics

Thesis Supervisor

Certified .....

Arnaud Costinot

Pentti J.K. Kouri Career Development Assistant Professor of

Economics

Thesis Supervisor

Certified by .....

..... Guido Lorenzoni

Associate Professor of Economics

Thesis Supervisor

Accepted by .....

Esther Duflo

Abdul Latif Jameel Professor of Poverty Alleviation and Development

Economics

Chairman, Departmental Committee on Graduate Studies



# Financial Distortions and the Distribution of Global

## Volatility

by

Maya Rachel Eden

Submitted to the Department of Economics  
on May 15, 2011, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Economics

### Abstract

In this thesis, I study the interactions between various aspects of the financial system and macroeconomic volatility in a globally integrated environment. In Chapter 1, I illustrate that an efficient allocation of liquidity across projects mitigates the economy's responsiveness to global liquidity supply shocks. Emerging economies in which the allocation of liquidity is distorted serve as a buffer zone that insulates developed economies from shocks to global liquidity supply. This suggests that, when functioning properly, the financial system in the developed world increases its stability by facilitating the efficient allocation of liquidity. However, I illustrate that in a global environment in which funding is cheap, the financial system will endogenously deteriorate and cease to carryout this role effectively. The conclusion is twofold: first, an efficient allocation of liquidity has a stabilizing effect on macroeconomic fluctuations. Second, in a low interest rate environment, the economy cannot rely on the financial system to maintain the capacity to implement an efficient allocation.

In Chapter 2, I suggest that intermediation need not be necessary in order to achieve an efficient allocation of liquidity; by setting an appropriately high tax on production or subsidy on unproductive savings, the government can manipulate the equilibrium prices of production inputs such that an efficient allocation of resources is achieved. Compared to the optimal policy benchmark, the equilibrium financial system absorbs too many productive resources. Further, the mere existence of a financial system induces unnecessary macroeconomic volatility in the form of liquidity shortages and surges in unemployment. I conclude that while the efficient allocation of liquidity is important both for the level of output and for output stability, financial intermediation is an inferior way to achieve it.

In Chapter 3, I study the distributional implications of allowing for the intermediation of liquidity from developed to emerging economies. Liquidity suppliers from developed economies extract rents from supplying liquidity to constrained entrepreneurs in emerging markets. Financial integration is therefore associated with a regressive transfer of surplus from emerging to developed economies. Further, as input prices in emerging economies appreciate following the inflow of liquidity, produc-

ers in emerging economies become increasingly reliant on foreign liquidity; a sudden reluctance of foreigners to supply liquidity results in a drop in output and consumption. Financial integration therefore not only decreases equilibrium consumption in emerging economies, but also increases the volatility of consumption due to shocks to external funding.

Thesis Supervisor: Ricardo J. Caballero  
Title: Ford International Professor of Economics

Thesis Supervisor: Arnaud Costinot  
Title: Pentti J.K. Kouri Career Development Assistant Professor of Economics

Thesis Supervisor: Guido Lorenzoni  
Title: Associate Professor of Economics



## Acknowledgments

Moses, it is said, had such a formative role in the development of the Israelites as a people in part because of his stutter: the ambiguity in his speech created fertile grounds for debate, interpretation, and scholarship. My primary adviser, Professor Ricardo J. Caballero, has had a similar role for me in my personal development as a scholar. Our discussions had a strong element of incompleteness: I was always left with more questions than answers. Through the process of reflecting on our conversations, pondering those questions and making them more and more precise, I managed to arrive at insights that, I feel, have deepened my understanding of economics. For this, as well as for his dedication, respect and kindness, I am very grateful.

I am also deeply indebted to my secondary advisers, Professor Arnaud Costinot and Professor Guido Lorenzoni. Professor Arnaud Costinot always took the time and care to go through my work thoroughly, and provided invaluable advice at all stages of writing this thesis. He helped me understand how to communicate effectively ideas of mine that did not always fit naturally into any standard framework and introduced me to the concept of log supermodularity, which is central to the first chapter of this thesis. Professor Guido Lorenzoni provided a fresh and often critical perspective on my work, and helped me both in understanding its appropriate relevance and in improving its exposition.

I am grateful for excellent training at the MIT department of economics and at the departments of mathematics and economics at the Hebrew University. For enriching and inspiring me, I thank my teachers at MIT, and especially Professor Daron Acemoglu, Professor George-Marios Angeletos, Professor Olivier Blanchard, Professor Dave Donaldson, Professor Bengt Holmstrom, Professor Francesco Giavazzi, Professor Roberto Rigobon and Professor Ivan Werning. I also thank my teachers at the Hebrew University, and especially Professor Sergiu Hart, Professor Ehud Hrushovski and Professor Motty Perry. I thank my fellow graduate students at MIT and especially Suman Basu, Joaquin Blaum, Fernando Duarte, Jonathan Goldberg, Luigi

Iovino, Daniel Keniston, Amir Kermani, Pablo Kurlat, Sahar Parsa, Michael Peters, Michael Powell, Jenny Simon and Simone Schaner, whose company enriched my thinking about economics and made my time at MIT more pleasant.

This thesis is dedicated to my grandparents, Dr. Shevach Eden and Zahava Eden, and to the memory of my grandparents, Dr. Israel Gregory Sohn and Bonnie Schooler Sohn. My maternal grandfather, Dr. Israel Gregory Sohn, inspired in me both a love for thinking and a love for knowledge. When I was a child he taught me how to play chess and took me on countless visits to the natural history museum where he worked. He was a micropaleontologist who relentlessly insisted on the superiority of scientific truths, and treated all else with skepticism. It is this skepticism that I believe served me most when writing this dissertation.

My maternal grandmother, Bonnie Schooler Sohn, passed away very shortly before I began my graduate studies at MIT. She was extremely proud of her family and traditions and maintained a tight network of her (very large) extended family. When I moved to Boston, I was immediately put in contact with various distant relatives, some of whom became close friends. Being a part of her family made this time in Boston a lot more enjoyable.

My paternal grandparents, Dr. Shevach Eden and Zahava Eden, both of them immensely gifted teachers, played a very active role in my education throughout my life, including in the completion of this PhD. They closely followed the progress of my graduate studies and took interest both in my research and my teaching. Conversations with them have always been uplifting, and leave me extremely happy.

My grandfather, Dr. Shevach Eden, always encouraged me to aspire to greatness, but also acknowledged more limited achievements as success. My grandmother, Zahava Eden, is one of the smartest and most confident women I know; her entrepreneurial attitude and creativity have served as an inspiration to me in many aspects of my life, especially when dealing with conflict.

I am grateful beyond words to my parents, Professor Ben Eden and Dr. Vivian Eden. From early childhood, they inspired in me a desire for scholarship and provided me with all the resources necessary to pursue it. I am not exaggerating when I say that

it is their love and support that got me through this PhD program. With nearly daily phone conversations, they provided me with laughter, advice, and most importantly, an invaluable sense of perspective.

I thank my brother, Mati Eden, for learning to welcome me on my irregular visits home, for the peace that he gives me and for the years of unconditional love. Parting from him was one of the most difficult things about pursuing this PhD, as I knew that he is unable to understand the reasons for my prolonged absences. He has learned to understand that I care for him even when I am away. I also thank my younger brothers, Michael Eden and Ittai Eden, for pleasant distractions and for providing a loving and fun home for me in the US.

Special thanks to Benny Bar-Lavi, a childhood friend who encouraged me to always be curious.

Finally, I would like to thank Matt Solomon, who stood by me through some of the harder times of completing this PhD. I am excited to be starting the next chapter of our lives together.

To my grandparents, Shevach Eden and Zahava Eden  
and  
to the memory of my grandparents, Israel Gregory Sohn and Bonnie Schooler Sohn

# Contents

<b>1</b>	<b>Financial Distortions and the Distribution of Global Volatility</b>	<b>17</b>
1.1	Introduction . . . . .	18
1.2	Related Literature . . . . .	22
1.3	A Simple Model of Financial Distortions . . . . .	25
1.3.1	Basic Environment . . . . .	25
1.3.2	The Financial System . . . . .	25
1.3.3	Microfoundations . . . . .	30
1.4	Closed Economy Equilibrium . . . . .	33
1.5	Volatility Divergence . . . . .	35
1.5.1	Integrated Equilibrium . . . . .	36
1.5.2	Small Open Economies . . . . .	40
1.6	The Endogenous Deterioration of the Financial System . . . . .	41
1.6.1	Endogenous Deterioration . . . . .	42
1.6.2	Normal Fluctuations . . . . .	43
1.6.3	Large Adverse Shocks . . . . .	45
1.7	Conclusion . . . . .	47
1.8	Appendix A: Figures . . . . .	53
1.9	Appendix B: Proofs . . . . .	55
1.9.1	Proof of Lemma 1 . . . . .	55
1.9.2	Proof of Lemma 2 . . . . .	55
1.9.3	Proof of Lemma 3 . . . . .	57
1.9.4	Proof of Proposition 1 . . . . .	58

1.9.5	Proof of Propositions 2 and 3 . . . . .	59
1.9.6	Proof of Proposition 4 . . . . .	63
1.9.7	Proof of Proposition 5 . . . . .	64
1.9.8	Proof of Lemma 4 . . . . .	66
1.9.9	Proof of Proposition 6 . . . . .	68
1.10	Appendix C: A Monetary Model of Liquidity Supply . . . . .	69
1.10.1	Liquidity Supply in the Closed Economy . . . . .	71
1.10.2	Liquidity Supply in the Integrated Economy . . . . .	75
<b>2</b>	<b>The Inefficiency of Financial Intermediation in General Equilibrium</b>	<b>79</b>
2.1	Introduction . . . . .	80
2.2	Benchmark: an Economy with no Financial Intermediation . . . . .	84
2.2.1	Optimal Policy . . . . .	87
2.3	An Economy with Financial Intermediation . . . . .	93
2.3.1	Is Financial Intermediation Welfare-Improving? . . . . .	94
2.3.2	Optimal Policy . . . . .	98
2.4	Generality . . . . .	99
2.5	Realistic Concerns . . . . .	100
2.5.1	Elastic Supply of Inputs . . . . .	101
2.5.2	Sticky Input Prices . . . . .	104
2.5.3	The Cost of Intermediation Compared to the Cost of Tax Col- lection . . . . .	106
2.6	Conclusion . . . . .	110
2.7	Appendix . . . . .	116
2.7.1	Proof of Lemma 10 . . . . .	116
2.7.2	Proof of Lemma 11 . . . . .	116
2.7.3	Proof of Proposition 8 . . . . .	117
<b>3</b>	<b>Global Imbalances in a Monetary Model of Liquidity Constraints</b>	<b>119</b>
3.1	Introduction . . . . .	119
3.2	Single Period Real Model . . . . .	123

3.2.1	The Unconstrained Benchmark . . . . .	123
3.2.2	The Constrained Economy . . . . .	125
3.3	A Dynamic Monetary Model of Liquidity Constraints . . . . .	130
3.4	Closed Economy Equilibrium . . . . .	133
3.5	Integrated Equilibrium . . . . .	136
3.6	Global Imbalances . . . . .	139
3.6.1	Optimal Policy . . . . .	142
3.7	Instability . . . . .	144
3.8	Conclusion . . . . .	146
3.9	Appendix . . . . .	150
3.9.1	Proof of Lemma 15 . . . . .	150
3.9.2	Proof of Proposition 9 . . . . .	151
3.9.3	Proof of Proposition 10 . . . . .	154
3.9.4	Proof of Lemma 17 . . . . .	156
3.9.5	Proof of Lemma 18 . . . . .	157





# List of Figures

1-1	The autarkic equilibria . . . . .	35
1-2	The integrated equilibrium . . . . .	39
1-3	The endogenous deterioration of the financial system . . . . .	44
1-4	Output volatility trend in developed economies . . . . .	53
1-5	Output volatility trend in emerging economies . . . . .	54
2-1	Tax costs . . . . .	107
3-1	The Chinese current account as a percent of gdp. . . . .	140
3-2	The Chinese current account as a percent of gdp: first differences on an annual frequency. . . . .	141



# List of Tables

2.1	Policy instruments . . . . .	88
2.2	Primitives . . . . .	89
2.3	Equilibrium objects . . . . .	89
2.4	Tax revenues and the budgets of collection agencies . . . . .	108



# Chapter 1

## Financial Distortions and the Distribution of Global Volatility

### Abstract

A generic feature of financial frictions, whatever their origins may be, is that they distort the allocation of funds to projects, causing some less productive projects to be funded while more productive projects are not. I formalize this idea by introducing a log supermodularity condition which requires that, at the margin, the difference in productivity between funded and unfunded projects is smaller in more distorted economies. Using this condition, I then revisit the relationship between financial distortions and macroeconomic volatility. My first set of results establishes that financial integration shifts the margin of adjustment to global liquidity shocks disproportionately to financially distorted regions, thereby providing a new and simple explanation for the divergent trends in the volatility of emerging and developed economies up to the recent crisis. My second set of results shows that a global environment in which liquidity is cheap is conducive to a deterioration of the financial system in the developed world. While cheap liquidity increases and stabilizes output in that region, it amplifies large adverse shocks.

**JEL Classification:** E44, F36, G01, G15

**Keywords:** financial distortions, log supermodularity, emerging market volatility, the Great Moderation, financial crises

## 1.1 Introduction

How does financial integration between emerging and developed economies affect the global distribution of output volatility? Why are shocks to external funding an important source of crises in emerging markets, but typically irrelevant in the developed world? What was the role (if any) of financial integration with emerging markets in setting the stage for the subprime crisis?

In this paper I develop a simple model of financial distortions to address these questions. I present two sets of results: first, I show that financial integration shifts the margin of adjustment to fluctuations in liquidity supply disproportionately from developed to emerging economies. In other words, shocks to external funding are an important source of crises in emerging markets *because* they are typically irrelevant in the developed world. Second, I show that a global environment in which liquidity is cheap is conducive to the deterioration of the financial system in developed economies. The deterioration of the financial system will amplify large contractions in liquidity supply; however, it will actually serve to increase and stabilize output during normal times. The model therefore generates volatility patterns that are consistent both with the amplification of the subprime crisis and with the euphoric pre-crisis environment, and suggests that financial integration with emerging economies may have led to endogenous structural changes in the financial system that set the stage for the subprime crisis.

The notion of financial distortions central to my simple model builds on the following idea. Regardless of whether distortions originate from collateral constraints, informational asymmetries or search frictions, financial distortions cause unproductive projects to be implemented before more productive projects. Because of this, the return to the marginal units of funding is relatively higher in more distorted economies. Stated formally, the marginal product of funds is log supermodular in funds and the level of distortion.

This simple condition has rich macroeconomic implications. An attractive feature of this approach is that it clearly disentangles the macroeconomic implications of

financial distortions from their micro source, emphasizing the generality of the former. More importantly, this approach isolates a particular feature of financial distortions that is relevant for understanding macroeconomic volatility. For example, this feature is tightly linked to the amplification of liquidity supply shocks: in more distorted economies, the marginal units of funding are relatively more important, as they are used to implement relatively more productive projects.

I embed this notion of financial distortions in a general equilibrium framework. Working capital is the single input of production and is determined by a fluctuating liquidity supply. In reduced form, a more distorted economy can be represented simply as an economy with a less steeply declining marginal product of working capital. The first result is that, holding the levels of financial distortions fixed, financial integration shifts the margin of adjustment to fluctuations in liquidity supply disproportionately to emerging economies, mitigating the effect of shocks to the supply of funding on the developed world. This result is closely tied to the assumption that financial distortions are more prevalent in emerging markets: compared to developed economies, implemented projects are less productive and unimplemented projects are more productive. A contraction in global liquidity supply will therefore lead to a disproportionate contraction in the amount of implemented projects, as relatively fewer previously-implemented projects are able to meet more stringent borrowing requirements. Similarly, an expansion in global liquidity supply will lead to a disproportionate expansion in the amount of implemented projects, as relatively more previously-unimplemented projects generate returns high enough to justify implementation. The equilibrium counterpart of emerging markets' vulnerability to shocks to external funding is a decline in liquidity-driven output fluctuations in developed economies. Emerging economies essentially serve as a buffer zone insulating developed economies from shocks to their own domestic liquidity supply.

This result proposes a novel link between the vulnerability of emerging economies to shocks to external funding, and the moderation of liquidity-driven output fluctuations in the developed world prior to the recent crisis. There is a vast literature emphasizing the role of international capital flows as a source of heightened volatility

in emerging markets<sup>1</sup>. Evidence suggests that emerging markets are more sensitive to shocks to the global supply of funds, and that, because of this, financial integration increases volatility in emerging market economies<sup>2</sup>. At the aggregate level, the last three decades, which have been characterized by rapid financial integration, have also been characterized by a divergence in the volatility levels of emerging and developed economies. Figures 1-4 and 1-5 plot the trends of the absolute value of the output gap (calculated based on a linear trend) in the developed and emerging market regions<sup>3</sup>. While the volatility of output has increased in the emerging market region, output volatility in the developed world has experienced a steep and steady decline, commonly referred to as the Great Moderation. As demonstrated by structural VAR analysis in Gali and Gambetti [2009], this decline seems to be attributed to the contribution of non-technology shocks, which, in the context of this model, can be interpreted as shocks to the supply of funding. The model in this paper suggests that these diverging trends can be explained as an equilibrium outcome of financial integration, holding the qualities of financial institutions fixed.

I extend the model to allow for the degree of financial distortions in the developed world to endogenously adjust to the financially integrated environment. Banks choose the level of financial distortions, where a more efficient allocation is associated with a higher cost. The second set of results relates to the idea that large flows of capital towards the developed world led to an endogenous deterioration of the financial system. In this model, the equilibrium level of financial distortions increases endogenously with financial integration. Intuitively, when liquidity is cheap most projects should be implemented anyway; the returns to sustaining institutions which enable

---

<sup>1</sup>See Broner and Ventura [2010], Uribe and Yue [2006], Fostel and Geanakoplos [2008], Neumeyer and Perri [2005], Fernandez-Villaverde et al. [2009] and Chang and Fernandez [2010].

<sup>2</sup>Examining a large panel of both emerging and developed markets that liberalized since the 1980s, Demirguc-Kunt and Detragiache [1999] document that the financial fragility induced by financial integration decreases with the quality of financial institutions. Stiglitz [2000] argues that financial liberalization in emerging markets is associated with more extreme shifts in capital flows, which make them more susceptible both to unsustainable boom and appreciation episodes, and to sudden stops. Further evidence of the fragility induced by financial integration in emerging markets is the drastic measures these economies take in order to insure against sudden stops and stabilize capital flows (for example, by accumulating large reserves, as documented by Caballero and Panageas [2005]).

<sup>3</sup>See Hakura [2007] for a more detailed discussion of volatility trends in emerging and developed economies.



differential implementation of projects based on their returns is low, so intermediaries as well as regulators may choose to refrain from doing so. In this model, these institutional changes have a mixed effect on output volatility: while they amplify large adverse shocks, they actually serve to increase and stabilize output during normal times, a feature consistent with the tranquil and prosperous pre-crisis environment. Intuitively, increased reliance on the implementation of low-quality projects makes the implementation of these projects worthwhile even in the presence of small shocks<sup>4</sup>. However, a large enough shock that deems the implementation of low-quality projects unprofitable entails an inevitable domino effect, as the implementation of productive projects must contract as well.

These findings suggest a role for financial integration in setting the stage for the subprime crisis. It has been suggested that low interest rates and high demand for US assets have led to the loosening of lending standards and various structural changes in the financial system that increased the relative importance of low-quality loans, such as increased securitization and increased reliance on securitized products in banks' balance sheets<sup>5</sup>. Interestingly, while these structural changes may have evidently amplified the shock to housing prices (as suggested by Brunnermeier [2009] and Gorton [2008]), they did not immediately lead to a crisis. Rather, capital flows to the developed world remained high and increasing as the financial system underwent these structural changes. In the context of this model, these puzzling dynamics can be rationalized by the mixed implications of a deterioration in the financial system with respect to small and large shocks.

The rest of the paper is organized as follows. In section 1.2 I discuss the related literature. In section 1.3 I introduce a novel reduced-form formulation of financial distortions, and present examples of microfoundations consistent with this formulation. In section 3.4 I embed the notion of financial distortions in a simple general equilibrium model, and characterize the closed economy equilibrium. In section 1.5 I

---

<sup>4</sup>This feature of the model is consistent with the findings of Keys et al. [2008], according to which increased securitization led to the loosening of lending standards and increased implementation of low-quality projects.

<sup>5</sup>See Brunnermeier [2009] and Gorton [2008].

consider the effects of financial integration between a distorted emerging market region and an undistorted developed market region. In section 1.6 I extend the model in section 1.5 to allow for a deterioration of the quality of the financial system in the developed world. I show that a decline in the price of liquidity may cause a deterioration in the quality of the financial system, and discuss the implications for output and output volatility. In section 3.8 I conclude.

## 1.2 Related Literature

In this section, I discuss the related literature and clarify the contribution of this paper relative to others.

The results concerning the divergence in volatility between emerging and developed economies are related to several strands of literature. The idea that financial distortions exacerbate output volatility appears prominently in the context of collateral constraints. Kiyotaki and Moore [1997], Fostel and Geanakoplos [2008] and Caballero and Krishnamurthy [2001] are important examples, the latter two with specific applications to emerging market economies. The general formulation in this paper suggests that the link between financial distortions and volatility is not unique to collateral constraints, and is common to many forms of distortions. The mechanism in this paper is closest to Kiyotaki and Moore [1997], as the amplification of shocks through the financial system is a result of a domestic inefficient allocation of resources. However, the role of external funding in explaining emerging market volatility is closest to Fostel and Geanakoplos [2008]. In Fostel and Geanakoplos [2008], a small group of constrained investors hold emerging market assets. Shocks to the liquidity in the hands of these investors evidently translate into movements in the liquidity supply for emerging market projects. The focus of this paper is complementary to the one in Fostel and Geanakoplos [2008]: while their starting point is that emerging market projects are funded by a small group of “residual” investors, this paper uses financial distortions to endogenize the fact that emerging market projects are “residual”.

Also related is the literature on volatility, development, and openness. The result that economies in earlier stages of development become more volatile with financial integration is in line with Obstfeld [1994], Greenwood and Jovanovic [1990], Acemoglu and Zilibotti [1997], and Koren and Tenreyro [2009]. The focus in these papers is on the changes in the sectoral composition induced by better consumption diversification opportunities. If financial integration allows for better diversification, investors may be more inclined to take on high-risk projects, potentially increasing the aggregate riskiness of the economy. This mechanism is very different from the mechanism discussed in this paper, which does not rely on changes in sectoral composition. While sectoral composition seems like a plausible source for some shocks, particularly in lower stages of development, it seems like an unlikely explanation for the heightened volatility of emerging markets associated with movements in external funds (such as sudden stops). Moreover, it provides a poor explanation for the common movements in emerging market economies, which have very different sectoral compositions.

Conceptually related is the work of Caballero et al. [2008], who study global imbalances as an outcome of financial integration in a world with heterogenous financial development. While the questions motivating these papers are different, both share the view that integration with a financially underdeveloped region is an important factor behind recent trends in the US economy. The work of Fogli and Perri [2006] links the Great Moderation with global imbalances by arguing that global imbalances are a natural artifact of the decline in volatility. In their model, the decline in volatility (relative to the rest of the world) reduces households' relative incentives to accumulate precautionary savings. This paper proposes an alternative link between financial integration and the Great Moderation. In this model, financial integration decreases volatility in the developed world and increases volatility in emerging markets; it could be that as a result, through the mechanism in Fogli and Perri [2006], the external balance of the developed markets deteriorates. This combined mechanism demonstrates an additional channel through which financial underdevelopment in emerging markets translates into global imbalances.

The results concerning the endogenous deterioration of the financial system in the

developed world are broadly related to two strands of literature. The idea that low interest rate environments are conducive to the deterioration of the financial system shares with the literature on bubbles. As shown in Tirole [1985], environments in which the interest rate is low are fertile grounds for the formation of bubbles. The formation of bubbles can be broadly viewed as a form of financial deterioration, as bubbles can be thought of as “low quality” investments with low fundamental value. Similar to the model presented here, the presence of bubbles may lead to an expansion in output (as in Farhi and Tirole [2010] and Martin and Ventura [2010]). However, the mechanisms through which a low interest rate leads to an expansion are different: in the bubble literature, bubbles serve as additional sources of liquidity, which enable more economic activity. In this paper, low interest rates are indicative of a situation in which the global supply of liquidity is already high; the expansion results from an endogenous decision of intermediaries to implement projects indiscriminately, increasing the total amount of projects implemented.

Similar to the literature on bubbles mentioned above, this paper takes the view that the root of the crisis is a sudden contraction in the supply of liquidity. The contraction in the supply of liquidity is not modeled here explicitly, and may be thought of as resulting from a burst of a bubble on an asset used for liquidity purposes (see Holmstrom [2008]). However, unlike the bubble-burst view of the crisis, in this model the burst of the bubble itself does not explain the full extent of the crisis. Rather, the crisis is amplified by the structural changes that the financial system underwent during the expansionary period. The emphasis on the role of financial frictions as an amplification mechanism of the crisis shares with Hall [2009], Gertler and Kiyotaki [2010], and others. The view closest to this paper is the one expressed in Brunnermeier [2009] and Gorton [2008]. These papers discuss mechanisms through which a low interest rate environment led to a decline in lending standards and institutional changes which evidently amplified the subprime crisis. The model presented in this paper may be seen as a simple formalization grouping these phenomena.

Methodologically, this paper is related to the literature emphasizing the role of supermodularity and log supermodularity conditions in various economic fields. Promi-

ment examples include, among others, Milgrom and Weber [1982] in auction theory; Bulow et al. [1985] in industrial organization; Jewitt [1987] and Athey [2002] in monotone comparative statics under uncertainty; Shimer and Smith [2000] in matching; and Costinot [2009] in international trade.

## 1.3 A Simple Model of Financial Distortions

In this section, I develop a notion of financial distortions, and show some microfoundations consistent with this formulation.

### 1.3.1 Basic Environment

Consider an economy in which there is a single consumption good, and a single input of production called *working capital*, denoted  $N$ . Working capital is funds used to hire workers, rent capital and buy intermediate inputs. Many of the results that follow will focus on the volatility of output attributed to the volatility of working capital; thus, working capital should be thought of as a variable affecting output through a fluctuating supply of funds<sup>6</sup>.

The fundamentals of the economy are given by an aggregate productivity level,  $A$ , and a set of *projects* indexed  $x \in (0, 1)$ . A project requires one unit of working capital to implement. If implemented, project  $x$  produces  $Ag(x) > 0$  units of output, where, without loss of generality,  $g$  is decreasing.

### 1.3.2 The Financial System

The role of the financial system is to allocate working capital to projects. I restrict attention to allocation schemes in which projects are implemented according to some *order*: the financial system organizes projects on a list. Given a supply of  $N < 1$  units of working capital, the first  $N$  elements on the list are to be implemented. Each unit of working capital may implement fractions of different projects, as long as the

---

<sup>6</sup>From an empirical standpoint, I prefer not to interpret  $N$  as physical capital, as physical capital does not fluctuate substantially at a business cycle frequency.

fractions sum up to 1. Equivalently, the fraction of implemented projects of each type is assumed to be weakly increasing in the level of working capital.

Formally, a financial system is represented by a function  $\sigma_\phi(N, x)$ , that determines the density of projects of type  $x$  implemented with the  $N$ th unit of working capital. At this point, the parameter  $\phi$  is simply an index of the financial system; later I will assume that it corresponds to the level of distortion. The function  $\sigma$  must satisfy the following condition, which guarantees that the total volume of projects implemented with each unit of working capital is 1, and that each project is fully implemented when the economy is satiated with working capital:

$$\int_0^1 \sigma_\phi(N, x) dx = \int_0^1 \sigma_\phi(N, x) dN = 1 \quad (1.1)$$

The marginal product of working capital is given by:

$$y(N, \phi) = \int_0^1 \sigma_\phi(N, x) Ag(x) dx \quad (1.2)$$

To see this, note that the above expression is a weighted average of productivities, where the weight on projects of type  $x$  is given by  $\sigma_\phi(N, x)$ , the density of projects of type  $x$  implemented with the  $N$ -th unit of working capital.

Aggregate output is given by the integral of the marginal product of working capital:

$$Y(N, \phi) = \int_0^N y(N', \phi) dN' \quad (1.3)$$

I restrict attention to allocation schemes that deliver a decreasing marginal product of working capital:

$$\frac{\partial y(N, \phi)}{\partial N} \leq 0 \quad (1.4)$$

I impose a partial ordering on the set of distortions. The notation  $\phi > \phi'$  indicates that the allocation scheme governed by  $\phi$  is more distorted than the allocation scheme governed by  $\phi'$ .

**Assumption 1** For all  $\phi > \phi'$  and  $N > N'$ :

$$1 \geq \frac{y(N, \phi)}{y(N', \phi)} > \frac{y(N, \phi')}{y(N', \phi')} \quad (1.5)$$

The assumption above states that the marginal product of working capital,  $y(N, \phi)$ , is log supermodular in  $N$  and  $\phi$ . If  $\ln y$  is differentiable with respect to  $N$ , this condition is equivalent to the statement that  $\frac{\partial \ln y(N, \phi)}{\partial N}$  is increasing in  $\phi$ . This turns out to be an important aspect of Assumption 1. In particular, if  $N$  is the level of working capital, Assumption 1 implies that for any small  $\epsilon > 0$ ,

$$1 \geq \frac{y(N + \epsilon, \phi)}{y(N - \epsilon, \phi)} > \frac{y(N + \epsilon, \phi')}{y(N - \epsilon, \phi')} \quad (1.6)$$

This inequality states that, around the marginal unit of working capital, the barely-implemented projects are more similar to the projects that just fall short from being implemented.

Intuitively, Assumption 1 means that better projects are implemented earlier on the list in less distorted economies; the decline in the relative quality of implemented projects is steeper the more efficient the allocation. The intuition is perhaps best understood when considering a simple economy in which there are only two projects: a good project which produces 3 if implemented, and a bad project which produces 1 if implemented:

$$g(\text{good}) = 3 \quad (1.7)$$

$$g(\text{bad}) = 1 \quad (1.8)$$

For whatever reason, the financial system errs with some probability  $\phi \in [0, \frac{1}{2}]$  and implements the bad project before the good project. The parameter  $\phi$  captures the level of distortion, as the probability of an error is higher in more distorted economies. The case  $\phi = 0$  corresponds to the efficient case. The case  $\phi = \frac{1}{2}$  corresponds to the case in which projects are implemented completely at random. In the general case,

the marginal product of working capital is:

$$y(1, \phi) = (1 - \phi) \cdot g(\text{good}) + \phi \cdot g(\text{bad}) = 3 - 2\phi \quad (1.9)$$

$$y(2, \phi) = (1 - \phi) \cdot g(\text{bad}) + \phi \cdot g(\text{good}) = 2\phi + 1 \quad (1.10)$$

Assumption 1 is satisfied since the following ratio is increasing in  $\phi$ :

$$\frac{y(2, \phi)}{y(1, \phi)} = \frac{2\phi + 1}{3 - 2\phi} \quad (1.11)$$

Assumption 1 is satisfied both because the return to the first unit of working capital is lower in more distorted economies, and because the return to the second unit of working capital is higher in more distorted economies.

More generally, Assumption 1 draws on the following principle. In efficient economies, only the most highly productive projects are implemented when funds are scarce. As funds become abundant, the economy runs out of highly productive projects, and less productive projects are implemented as well. In distorted economies, the order in which projects are implemented is inefficient. Scarce funds are used to implement an inferior set of projects; some unproductive projects are implemented, while some highly productive projects are not. This suggests that the quality of implemented projects is closer to the quality of unimplemented projects. At the margin, the expected ratio of the quality of just-implemented projects and just-not-implemented projects is lower in more distorted economies, reflecting the idea that the implementation outcome is more affected by factors unrelated to productivity.

Some notes are in order regarding the limitations of this approach. Assuming that projects are implemented according to some order is somewhat restrictive, as not all microfoundations of distortions take this form. This restriction rules out distortions in which some projects are implemented when resources are scarce, but not implemented when resources are abundant. However, in its reduced form, Assumption 1 will continue to hold as long as an increase in the amount of funding is associated



with an increase in the efficiency of the allocation. This is in the spirit of Kiyotaki and Moore [1997], and consistent with other models of financial distortions, such as Rajan [1994]. Intuitively, in these models the returns to the marginal units of funding have a relatively high return because they increase the productivity of the inframarginal units of funding.

Even under the restriction that projects are implemented according to some order, there are alternative ways to rank distortions that do not necessarily imply compliance with Assumption 1. A particularly appealing way to rank distortions is according to the productivity loss that they induce: a more distorted economy can be defined naturally as one in which output is lower for any given level of working capital. Assumption 1 is more restrictive than compliance with this property. However, there is some comfort in the fact that Assumption 1 guarantees that there is a productivity loss induced by higher financial distortions:

**Lemma 1** *The level of output  $Y(N, \phi)$ , is decreasing in the distortion parameter  $\phi$ .*

The proof is omitted from the text and, together with other omitted proofs, can be found in the appendix. The converse of Lemma 1 is not necessarily true: requiring that a higher distortion parameter induces a productivity loss need not imply compliance with Assumption 1. Intuitively, the added restriction in Assumption 1 is that in more distorted economies, the implementation outcome depends less on productivity at *any* level of working capital. It is possible to construct examples of distortions in which productivity matters less for the implementation outcome at some levels of working capital, but matters more at other levels of working capital<sup>7</sup>.

While Assumption 1 does not include all possible models of distortions, it points at a property that is fairly general and common to many microfoundations, as will be demonstrated in the section that follows. The particular property emphasized by Assumption 1 turns out to have far-reaching implications for macroeconomic volatility,

---

<sup>7</sup>Consider the following simple example, in which an elite set of projects indexed  $x \in (0, \bar{x})$  is implemented according to the efficient order, and the rest of the projects are implemented randomly. Of course, compared to the efficient allocation, there is a productivity loss induced by this distortion. However, Assumption 1 is violated, as at  $N = \bar{x}$ , there is a larger difference in the average productivities of projects implemented at the margin and projects just-not-implemented.

which may not be true under less restrictive conditions.

### 1.3.3 Microfoundations

In this section I present three microfoundations of distortions. The first microfoundation will be used to illustrate the results throughout the paper.

**Asymmetric information (random allocation).** Consider a stark model in which there are two types of economies: efficient and distorted. In efficient economies, the financial system employs a technology that allows projects to be screened according to their type. Projects are implemented in the efficient order. In distorted economies, the financial system lacks the technology to distinguish between projects. As project owners receive large private benefits from implementing their projects, the only equilibrium is a pooling equilibrium in which projects are implemented indiscriminately. As a result, projects are implemented completely at random: the order in which projects are implemented is completely unrelated to their productivity.

Formally, in the efficient economy, projects are implemented in an order decreasing in their returns; given any  $N$  units of working capital, the first  $N$  projects will be implemented. Output is given by the aggregation of the productivities of all implemented projects:

$$Y(N, \text{efficient}) = \int_0^N Ag(x)dx \tag{1.12}$$

The marginal product of working capital is therefore given by:

$$y(N, \text{efficient}) = Ag(N) \tag{1.13}$$

Note that, in the efficient allocation, the function  $\sigma$  is a Dirac measure, implementing the entire project indexed  $x$  with the  $x$ -th unit of working capital. The log of the marginal product of working capital is decreasing, as the productivity of implemented projects declines with working capital.

In contrast, in the distorted economy, a random set of projects is implemented. The function  $\sigma(N, x)$  is equal to 1 for all  $x$  and  $N$ , capturing the fact that each unit

of working capital implements a random set of projects. Output is given by:

$$Y(N, \text{distorted}) = N \int_0^1 1 \cdot Ag(x)dx \quad (1.14)$$

The marginal product of working capital is therefore given by:

$$y(N, \text{distorted}) = \int_0^1 1 \cdot Ag(x)dx \quad (1.15)$$

Since the log of the marginal product is decreasing in  $N$  in the efficient economy and constant in the random allocation, the random allocation is *more distorted* than the efficient economy in accordance with Assumption 1<sup>8</sup>.

**Collateral constraints.** Consider a model in which projects are characterized by their type,  $x$ , and by their collateral type,  $b$ . Assume that collateral types ( $b$ ) are distributed independently from the project type ( $x$ ). Conditional on any project type, collateral is uniformly distributed on  $[0, \phi]$ . The collateral level of the project is given by the following decreasing function:

$$\kappa(b) = g(\min\{b, 1\}) \quad (1.16)$$

That is, collateral is decreasing in  $b$ , and is bounded from below by the productivity of the least productive project.

A project with features  $(x, b)$  is implemented if and only if both its return and its collateral level exceed the price of working capital:

$$g(x) \geq r \quad (1.17)$$

$$\kappa(b) \geq r \quad (1.18)$$

---

<sup>8</sup>Alternatively, rather than assuming that the random allocation is a result of this form of asymmetric information, this allocation may result from institutional arrangements under which the incentives to implement projects are detached from their returns. This may be the case, for example, in the presence of corruption or government expropriation.

The mismatch between the quality of projects and their collateral changes the order in which projects are implemented. Note that a higher  $\phi$  implies that the aggregate collateral is lower. It is easy to see that  $\phi \rightarrow 0$  corresponds to the efficient allocation, as all projects have sufficient collateral to be implemented when their return is high enough. The case  $\phi \rightarrow \infty$  corresponds to an extremely distorted economy, in which essentially all projects are collateral constrained and can be implemented only when the market return is equal to the productivity of the least productive project; in this case, the marginal product of working capital is close to constant, as projects are implemented essentially in a random order (as in the previous example).

**Lemma 2** *Assume that  $g(x) = \alpha x^{-(1-\alpha)}$ . In this model of collateral constraints, Assumption 1 is satisfied.*

**Search frictions.** Consider a model in which economies differ in the search technology available to their financial systems. In each economy there are  $\phi$  local banks indexed  $i = 0, \dots, \phi - 1$ . Projects owners are unaware of their project's type until right before production decisions must be made. Local banks are modeled as risk sharing arrangements among project owners. Each local bank shares risk among  $\frac{1}{\phi}$  project owners. The bank indexed  $i$  is a risk sharing arrangement between the owners of projects indexed  $x \in (\frac{i}{\phi}, \frac{i+1}{\phi}]$  (for  $i = \phi - 1$ , the segment of projects is the open set  $(\frac{\phi-1}{\phi}, 1)$ ). The process of distributing funds from households to projects is as follows:

1. Households supply funds to a large savings bank. The savings bank has direct access to projects (but is unaffiliated with the local banks, and is unaware of projects' types).
2. The savings bank distributes funds randomly among projects owners.
3. Project owners handover their funds to their local bank. When the productivity of projects is revealed, the local bank uses the funds in its hands to implement the best set of projects among those owned by the bank's members.

The parameter  $\phi$  can be thought of as a measure inversely related to the integration of the domestic financial system. A high  $\phi$  captures a situation in which there are some

banks with access to highly productive projects that lack liquidity to implement them, and some banks with high liquidity without access to good projects. The case  $\phi = 1$  corresponds to the efficient allocation: there is only one bank in charge of allocating the entire supply of funds, and the bank has access to the entire set of projects. The optimal set of projects is therefore implemented. The case  $\phi \rightarrow \infty$  corresponds to the random allocation example: project owners are essentially in autarky, as they must use their own funds to implement their own project. The implementation of projects is therefore random, as it does not depend on the level of productivity.

**Lemma 3** *Assume that  $g(x) = e^{-x}$ . In this model of search frictions, Assumption 1 is satisfied.*

## 1.4 Closed Economy Equilibrium

The characterization of the closed economy equilibrium is useful in two ways. First, it provides a benchmark for comparison with the open economy. Comparing the closed economy equilibrium with the open economy equilibrium will be useful for understanding the implications of financial integration. Second, the closed economy equilibrium isolates a particular mechanism through which financial distortions amplify output volatility, which will be important both in section 1.5 and in section 1.6.

Households inelastically supply  $Q$  units of *liquidity* to the financial system, in exchange for future returns. In this model, liquidity is defined simply as the supply side of working capital. The market clearing condition is:

$$N = Q \tag{1.19}$$

After production takes place, the financial system repays households at a rate of  $r$  units of output per unit of liquidity. I assume that  $r$  is equated with the marginal product of working capital. This can be thought of as a result of a competitive banking system in which banks compete for liquidity supply from households.

There are two sources of volatility: shocks to the domestic technology level  $A$ , and shocks to the supply of liquidity,  $Q$ . I assume that these shocks are independent. Appendix 1.10 presents a model of liquidity supply in which this assumption is satisfied. In that model, liquidity supply fluctuations are driven by shocks to the money supply and shocks to investor's risk aversion.

It is straightforward to show that, in the closed economy, the sensitivity of output to TFP shocks is unrelated to the degree of financial distortions. However, more distorted economies are more sensitive to shocks to liquidity supply:

**Proposition 1** *For any given processes of  $A$  and  $Q$ ,*

1. *Output is more volatile in more distorted economies:*

$$\frac{\partial \ln Y(Q, \phi)}{\partial \ln Q} \text{ is increasing in } \phi \quad (1.20)$$

2. *Average productivity is more sensitive to liquidity supply shocks in less distorted economies:*

$$\frac{\partial \frac{Y(Q, \phi)}{Q}}{\partial \ln Q} \text{ is increasing in } \phi \quad (1.21)$$

Proposition 1 follows from the property of financial distortions emphasized in Assumption 1. In relatively more efficient economies, the decline in the average productivity of projects mitigates the effect of an increase in working capital. In distorted economies, a suboptimal set of projects is implemented, so some low-yield projects are implemented before higher-yield projects. This implies two things. First, some projects that should have been implemented are not implemented; their implementation takes place only when the supply of working capital is higher, intensifying the returns to working capital. Second, some projects that should not have been implemented are implemented, lowering the average product of working capital in the economy. These two facts put together imply that the ratio of the marginal product of working capital and the average product of working capital is higher in distorted economies. In other words, liquidity abundance increases the efficiency of distorted economies relative to efficient economies both by alleviating the inefficiency caused by

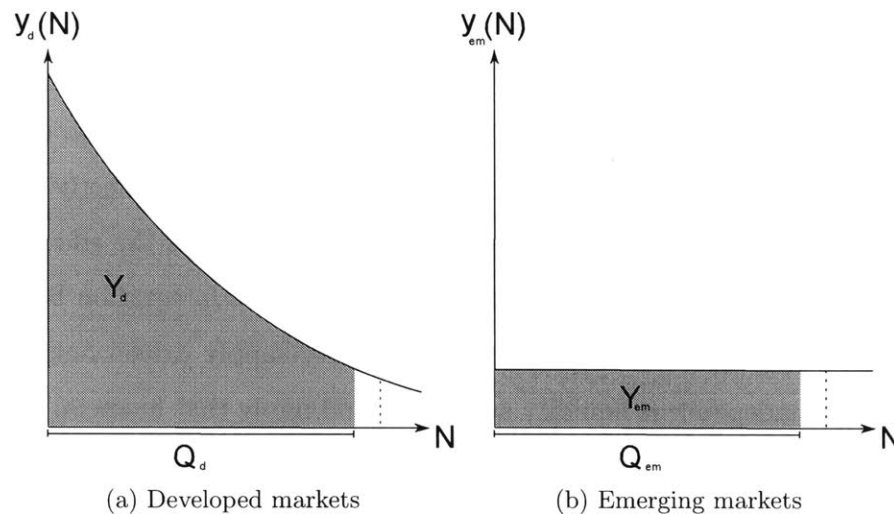


Figure 1-1: The autarkic equilibria

implementing a suboptimal set of projects and by allowing for the implementation of higher-yield projects. Using the random allocation example, figure 1-1 illustrates the properties of the closed economy equilibria in a distorted “emerging market” economy and an efficient “developed market” economy.

## 1.5 Volatility Divergence

In this section I present a set of results concerning the divergence of volatility between emerging and developed economies following financial integration. I first present the globally integrated equilibrium, and then discuss equilibrium implications for small open economies.

The setup of the model is as follows. There are two regions of equal size: an emerging market region (*em*) and a developed market region (*d*). For the time being, I assume that the only difference between emerging and developed economies is that emerging economies are more distorted than developed economies:

$$\phi_{em} > \phi_d \tag{1.22}$$

I assume that  $\text{var}(\ln(A_d)) = \text{var}(\ln(A_{em}))$ <sup>9</sup> and that liquidity supply  $Q_i$  is independent from both foreign and domestic technology,  $A_i$  and  $A_j$ .

In order to isolate the effects of financial heterogeneity on the global equilibrium environment, it is convenient to assume that  $Q_d$  and  $Q_{em}$  are perfectly correlated. Assuming that liquidity supplies are perfectly correlated isolates the effects of financial heterogeneity because, under this assumption, financial integration between two *identical* economies would have no effect on liquidity-supply driven output volatility. In contrast, independent liquidity supplies would imply that financial integration between two identical regions has a moderating effect on output in both regions, as shocks to liquidity supply are shared across regions. Replacing the assumption that  $Q_d$  and  $Q_{em}$  are perfectly correlated with the assumption that they are independent would therefore decrease volatility in both regions; however, the result that volatility induced by financial integration would be relatively higher for emerging economies would still hold.

### 1.5.1 Integrated Equilibrium

Assume that liquidity can move freely across regions. The global equilibrium is characterized by two equations:

$$N_{em} + N_d = Q_{em} + Q_d = Q_w \quad (1.23)$$

$$y_{em}(N_{em}, \phi_{em}) = y_d(N_d, \phi_d) = r \quad (1.24)$$

The first equation is a market clearing condition, stating that the total amount of working capital,  $N_{em} + N_d$ , must be equal the global supply of liquidity,  $Q_w = Q_{em} + Q_d$ . The second condition is the optimality condition of the financial system, requiring that there is no gain from reallocating liquidity from one region to another.

I denote autarkic values with superscript  $a$  ( $N^a$ ,  $Y^a$ , etc). Denote by  $\Delta$  the absolute value of the *average* change in working capital levels induced by financial

---

<sup>9</sup>The results trivially generalize to  $\text{var}(\ln(A_d)) < \text{var}(\ln(A_{em}))$ .



integration:

$$\Delta = |E(N_{em}) - E(N_{em}^a)| = |E(N_d) - E(N_d^a)| \quad (1.25)$$

The value of  $\Delta$  is determined in equilibrium as a function of the average supply of liquidity, the financial distortions, and the relative productivities  $\frac{A_d}{A_{em}}$ . In particular, there is always a value of  $\frac{A_d}{A_{em}}$  for which  $\Delta = 0$ . For the purpose of this exercise, I assume that  $\Delta$  is small. The importance of this assumption is in assuring that the sensitivity of output to working capital remains similar to its autarkic level. If this assumption is violated, the implications of financial integration on macroeconomic volatility depend more specifically on how the sensitivity of output with respect to working capital changes with the level of working capital.

The main result is stated in the proposition below:

**Proposition 2** *For  $\Delta$  sufficiently small, financial integration exacerbates the volatility differences between emerging and developed markets:*

$$\text{var}(\ln N_{em}) - \text{var}(\ln N_d) > \text{var}(\ln N_{em}^a) - \text{var}(\ln N_d^a) = 0 \quad (1.26)$$

$$\text{var}(\ln Y_{em}) - \text{var}(\ln Y_d) > \text{var}(\ln Y_{em}^a) - \text{var}(\ln Y_d^a) > 0 \quad (1.27)$$

Financial integration leads to a divergence in volatility levels for two reasons<sup>10</sup>. First, financial integration is associated with a new source of fluctuations in working capital, which is shocks to the relative productivity of emerging and developed economies ( $\frac{A_{em}}{A_d}$ ). These shocks lead to a substitution of working capital across regions. Since, by proposition 1, financially-distorted emerging markets are more sensitive to fluctuations in working capital than developed markets, this works towards exacerbating the differences in output volatility across regions.

Second, it turns out that shocks to the global supply of liquidity adjust disproportionately through changes in the supply of working capital to emerging markets.

---

<sup>10</sup>Proposition 2 easily generalizes to an equilibrium with many countries with different levels of distortions. In an integrated equilibrium, the sensitivity of each economy to shocks to the global supply of liquidity or to TFP will be an increasing function of its level of financial distortions. Differences in volatility levels will magnify upon financial integration.

This equilibrium property is closely related to the feature of financial distortions emphasized by Assumption 1, and particularly to its interpretation in equation 1.6. In developed markets, projects are implemented in an order decreasing in their returns: most implemented projects generate returns which well exceed  $r$ , and most of the unimplemented projects generate returns which are well below  $r$ . Fluctuations in the implementation threshold therefore have a relatively small impact on the amount of projects implemented in developed economies. In contrast, in emerging markets, the order in which projects are implemented is more arbitrary; this implies that the return generated by the next unit of working capital is similar to the return generated by the previous unit. The same fluctuations in  $r$  therefore induce larger fluctuations in the amount of implemented projects. Figure 1-2 illustrates this equilibrium property in the random allocation example<sup>11</sup>.

Generically, abstracting from productivity shocks, financial integration leads to less variation in  $r$  from the developed market's perspective, and more variation in  $r$  from the emerging market's perspective. Consequently, there is a divergence in liquidity-driven output volatility between emerging and developed economies.

The direction in which volatility levels change following financial integration is potentially different in emerging and developed economies. In emerging economies, the volatility of working capital necessarily increases upon integration. This is because working capital becomes vulnerable to two new shocks: shocks to foreign liquidity supply ( $Q_d$ ) and shocks to relative technology levels ( $\frac{A_{em}}{A_d}$ ). In developed economies, working capital also becomes vulnerable to shocks to relative technology, which works towards increasing volatility. However, working capital becomes less sensitive to shocks to domestic liquidity supply ( $Q_d$ ), since shocks to liquidity supply adjust primarily through movements in the working capital supplied to emerging markets. The net effect on volatility depends on the relative importance of technology shocks and liquidity shocks. Specifically, if movements in the level of working capital result primarily from variation in the supply of liquidity, financial integration will decrease the

---

<sup>11</sup>In the random allocation example, the equilibrium level of  $r$  does not change following a shock to global liquidity supply. Rather, the shock is adjusted entirely through changes in the quantity of working capital supplied to the emerging market region.

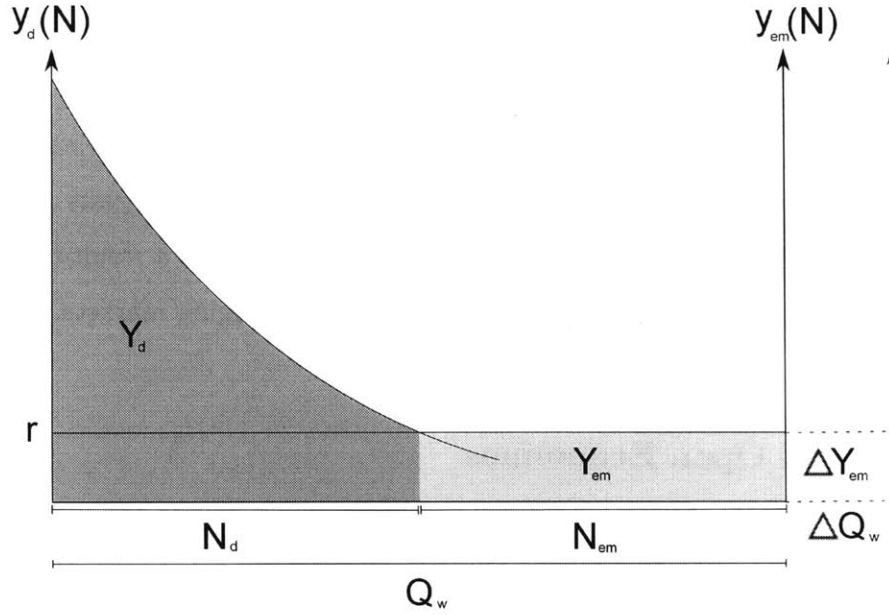


Figure 1-2: The integrated equilibrium

volatility of output in developed economies.

Under stronger assumptions regarding the relative importance of liquidity and technology shocks, it is possible to obtain stronger results concerning the direction in which volatility levels change following financial integration:

**Proposition 3** *For  $\Delta$  sufficiently small,*

1. *Financial integration increases the volatility of working capital and the volatility of output in emerging markets:*

$$\text{var}(\ln N_{em}) > \text{var}(\ln N_{em}^a) \quad (1.28)$$

$$\text{var}(\ln Y_{em}) > \text{var}(\ln Y_{em}^a) \quad (1.29)$$

2. *If the variance of liquidity supply is sufficiently large compared to the variance of relative TFP, financial integration decreases the volatility of working capital and the volatility of output in developed markets:*

$$\text{var}(\ln N_d) < \text{var}(\ln N_d^a) \quad (1.30)$$

$$\text{var}(\ln Y_d) < \text{var}(\ln Y_d^a) \tag{1.31}$$

These results provide some insight into the distinct behavior of emerging and developed markets following globalization. The theory above suggests that the divergence in liquidity-driven output fluctuations may have been a result of developed markets effectively exporting their liquidity shocks to emerging markets.

### 1.5.2 Small Open Economies

Given that much of the literature on the excess volatility of emerging markets has focused on small open economies, it is useful to study small open economies in the context of this global equilibrium environment.

I assume that each region is composed of a continuum of small open economies, identical within regions. Small open economies are subject to idiosyncratic productivity shocks, as well as to exogenous shocks to  $r$  which result from changes in the global liquidity supply.

The first thing to note is that this model naturally implies comovements in emerging market economies. Compared to their developed counterparts, small open emerging markets are more severely affected by shocks to the global liquidity supply. The importance of global liquidity supply as a source of emerging market fluctuations naturally implies common movements in emerging market output levels. Similar to Fostel and Geanakoplos [2008], comovements in emerging market economies result from a common sensitivity to an external supply of funding.

Second, it is interesting to note that both in emerging and in developed markets, working capital responds similarly to idiosyncratic productivity shocks and to shocks to the price of liquidity. This suggests a link between the heightened sensitivity of emerging markets to interest rate shocks (as in Neumeyer and Perri [2005] and Uribe and Yue [2006]) and the amplification of shocks to productivity (as in Caballero et al. [2005]). To see this link, note that in a small open economy, the level of working capital is pinned down by a single indifference condition, equating the marginal product of

working capital with the world rate of return:

$$y(N, \phi) = \int_0^1 \sigma_\phi(N, x) A g(x) dx = r \Rightarrow \int_0^1 \sigma_\phi(N, x) g(x) dx = \frac{r}{A} \quad (1.32)$$

From the formulation above, it is easy to see that working capital is affected similarly by shocks to  $r$  and shocks to  $A$ . Essentially, in this model, the responsiveness of working capital to either type of shock captures the density of projects which are implemented at the margin and collectively yield a return equal exactly to  $r$ . A small shock to the returns of these projects will shift them above or below the implementation threshold; similarly, small shocks to  $r$  will determine whether or not the projects at the margin generate a return which justifies implementation. The result that small open emerging markets are relatively more sensitive to both shocks is closely tied to Assumption 1, as it guarantees that the density of projects implemented at the margin is higher.

The model presented in this section suggests the following conclusions regarding the role of financial institutions in determining the effects of financial integration on output volatility. Poor financial institutions in emerging markets exacerbate their sensitivity to working capital, as well as increase the volatility of working capital supplied to that region. At times in which financial institutions are intact in developed markets, their superior ability to implement projects differentially serves both to stabilize equilibrium working capital levels and to mitigate the effects of fluctuations in working capital on output.

## 1.6 The Endogenous Deterioration of the Financial System

In section 1.5 it was assumed throughout that the quality of financial institutions remains fixed upon financial integration. While this may be a valid short-run assumption, the recent subprime crisis comes as a reminder that the quality of financial

institutions may evolve with changing circumstances.

In section 1.6.1 I consider a model in which the financial system in the developed world deteriorates endogenously following financial integration. In section 1.6.2 I discuss the implications of the deterioration during “normal times”. I show that in the absence of large shocks, the deterioration of the financial system actually increases and stabilizes output in the developed world. In section 1.6.3 I show that the deterioration in the financial system amplifies large adverse shocks.

### 1.6.1 Endogenous Deterioration

I extend the setup to allow for an endogenous adjustment in the quality of the financial system in the developed world. Banks can choose the level of financial distortions out of some finite set. There is a cost  $\lambda(\phi)$  associated with choosing the level of distortions  $\phi$ , where  $\lambda(\cdot)$  is decreasing. This cost should be thought of as the cost of sorting projects and overcoming other obstacles which stand in the way of an efficient allocation. For simplicity, I assume that banks consider only the mean price of liquidity when choosing the level of financial distortions, and do not take into account any uncertainty.

Given a price of liquidity  $r$ , the banks choose the level of working capital,  $N_d$ , and the level of distortion,  $\phi_d$ , to maximize profits. The bank’s profits are given by:

$$\pi(r) = \max_{\phi_d, N_d} \{ \pi(r, N_d, \phi_d) \} = \tag{1.33}$$

$$\max_{\phi_d, N_d} \int_0^{N_d} y(N', \phi_d) dN' - rN_d - \lambda(\phi_d) \tag{1.34}$$

Note that for  $\frac{A_d}{A_{em}}$  sufficiently large, financial integration will be associated with a decline in  $r$  from the developed market’s perspective.

**Proposition 4** *For  $Q$  sufficiently large, a decline in  $r$  will lead to an endogenous deterioration of the financial system.*

It is straightforward to show that the objective of banks in this model coincides with the objective of a social planner trying to maximize domestic output minus the

costs of differentiation (under assumption that the cost of liquidity is given by  $r$ ). The deterioration in the financial system can therefore be interpreted more broadly as an outcome of *endogenous* lax regulation.

In the context of the recent crisis, the above proposition formalizes the popular claim according to which financial integration increased the equilibrium level of financial distortions by lowering the price of liquidity from the developed market's perspective<sup>12</sup>. Intuitively, if liquidity is sufficiently cheap so that nearly all projects are implemented anyway, the benefits of differentiating between projects may not be worth the cost<sup>13</sup>. Broadly interpreted, the financial system will gravitate towards various institutional arrangements that tie the implementation of high quality projects with the implementation of low quality projects.

## 1.6.2 Normal Fluctuations

What are the implications of the deterioration of the financial system on output and on output volatility? Under autarky, the implications would be fairly intuitive: a weakening of the financial system would decrease output and increase subsequent output volatility<sup>14</sup>. However, these intuitive results breakdown under financial integration. The following proposition states that the deterioration in the quality of the financial system in the developed world will increase working capital in that region,

---

<sup>12</sup>A note is in order regarding the applicability of these results to emerging market economies. The analysis in this section relies heavily on the assumption that financial integration is associated with a decline in the price of liquidity from the domestic perspective. The results are therefore not applicable for emerging markets. For emerging markets, the mirror image is the relevant one: financial integration is associated with an increase in  $r$ , potentially leading to an endogenous increase the the quality of financial institutions. The characterization of the long run general equilibrium environment, in which financial institutions are allowed to adjust in both emerging and developed economies, is beyond the scope of this paper. I leave this interesting issue for future research.

<sup>13</sup>It is worth noting that, in general, the relationship between  $r$  and the equilibrium choice of financial distortions is non-monotonic. At sufficiently low levels of  $r$ , a decrease in  $r$  will decrease the benefits of investing in high-quality financial institutions. However, it is easy to show that, at high levels of  $r$ , a small drop in the price of liquidity will have the opposite effect. This is because there is a scale effect: the investment in the ability to differentiate is only worthwhile if production is sufficiently high.

<sup>14</sup>To see this, note that by Lemma 1, holding working capital fixed, output is higher when financial institutions are intact (because a superior set of projects are implemented). Thus, a weakening of the financial system constitutes an adverse shock to output. The fact that subsequent output volatility increases is immediate from the comparison between closed emerging and developed economies in section 3.4.

and may increase and stabilize output as well:

**Proposition 5** *In a financially integrated equilibrium:*

1. *An endogenous weakening of the financial system in the developed market will **increase** working capital in that region.*
2. *For  $Q_w$  and  $\frac{A_d}{A_{em}}$  sufficiently large, a weakening of the financial system in the developed world will increase and stabilize output.*

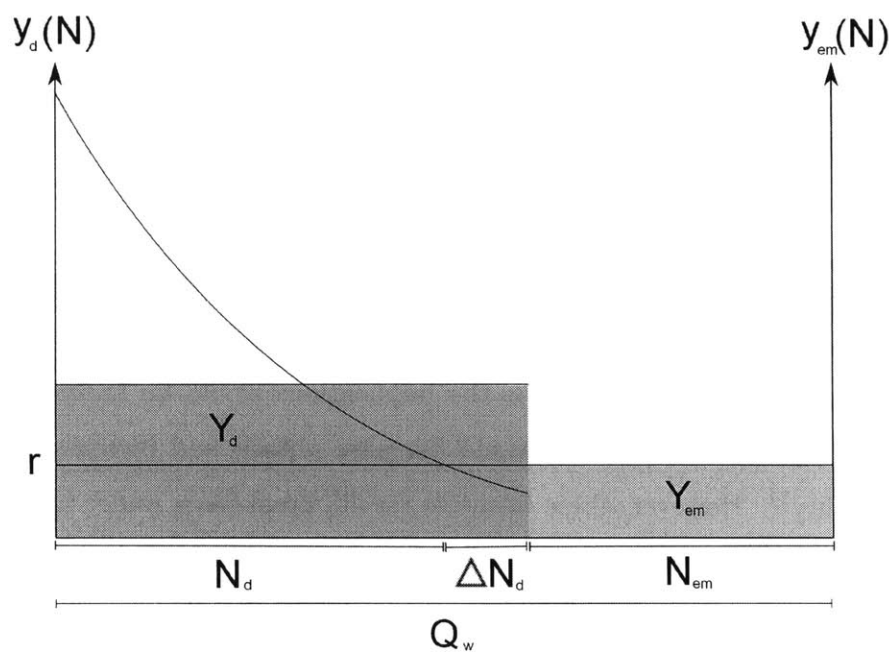


Figure 1-3: The endogenous deterioration of the financial system

Proposition 5 builds on a comparative static result that shows that an endogenous deterioration of the financial system is possible in equilibrium only if it increases the marginal product of working capital. The weakening of financial institutions is therefore associated with an additional flow of working capital towards developed economies. If this flow is sufficiently large, domestic output will rise. To understand why volatility declines, note that the distortion ties the implementation of high-quality projects to the implementation of low-quality projects. Thus, many “marginal” projects become inframarginal, as their discontinuation would necessitate the discontinuation of some high-quality projects.



Proposition 5 can shed light on the seemingly “irrational” behavior that led up to the subprime crisis, in which many bad loans were given to subprime borrowers, and, while the financial system behaved “irresponsibly”, the demand for US assets seemed only to increase.

Finally, does the fact that output in the developed world increases with the deterioration of its financial system mean that such a deterioration is “good”? From a global perspective, no. The generic adverse effect of a deterioration in the financial system is that it decreases world output, both under autarky and under financial integration:

**Lemma 4** *A weakening of the financial system in the developed world decreases world output.*

This lemma is immediate from the fact that intact financial institutions allow for a differential implementation of projects based on their returns; the set of projects implemented when differentiation is not possible is necessarily inferior to the set of projects implemented when some differentiation is possible.

### 1.6.3 Large Adverse Shocks

In the global environment described in this paper, the deterioration of the financial system in the developed world serves to increase and stabilize output in that region during normal times. However, while the response to small shocks is mitigated, the output response to large negative shocks is amplified:

**Proposition 6** *In the developed world, the weakening of the financial system amplifies the output response to large adverse shocks to  $Q_w$  or  $\frac{A_d}{A_{em}}$ .*

Intuitively, cheap liquidity leads to structural changes in the financial system that disable the separation of high-quality projects from low-quality projects. During normal times, this increases the amount of low-quality projects being implemented, and output increases as a result. However, sufficiently large contractions in liquidity

are amplified by the inability to separate the discontinuation of low-quality projects from the discontinuation of high-quality projects.

This result suggests that the deterioration in the quality of financial institutions preceding the sub-prime crisis may have indeed precipitated it by creating an amplification mechanism for large adverse shocks. This amplification mechanism is consistent with many inefficiencies that seem important for understanding the extent of the crisis. A straightforward interpretation is the mortgage market itself. The creation of mortgage backed securities enabled the pooling of idiosyncratic risk of subprime loans. Once housing prices declined, issuing new subprime loans became difficult, perhaps in part because the process of issuing subprime loans did not allow for differentiation between relatively promising borrowers and relatively unpromising borrowers. A more subtle interpretation is the balance sheet effect (as in Brunnermeier [2009]). The heavy reliance of banks' balance sheets on mortgage backed securities forced them to disengage from productive lending activities once the subprime crisis hit. This can be viewed as an additional mechanism that ties the implementation of productive projects to the implementation of unproductive projects.

**Corollary 1** *Financial integration may lead to the amplification of large adverse shocks in the developed world.*

Corollary 12 is immediate from the analysis in this section: by Proposition 4, financial integration leads to the deterioration of the financial system. By Proposition 6, the deterioration in the financial system amplifies the output response to large adverse shocks.

Note that this analysis also suggests that the subprime crisis cannot be explained solely in terms of a breakdown in the financial system; rather, a complete explanation would require either an additional large shock to TFP or to liquidity supply.

The results in this section present a modification to the view presented in section 1.5 according to which financial integration mitigates output fluctuations in the developed world. If financial integration with a distorted emerging market region is coupled with a decline in the price of liquidity (as suggested by Caballero et al.

[2008]), the quality of financial institutions may adjust downwards in accordance with Proposition 4. As a result, the sensitivity of output with respect to small shocks will decrease, but large adverse shocks will be amplified.

## 1.7 Conclusion

This paper studies the effects of financial distortions on the global equilibrium environment. I present a reduced form formulation of financial distortions according to which the marginal product of working capital is log supermodular in working capital and the level of distortion. This formulation is consistent with a class of microfoundations in which the distortion changes the order in which projects are implemented, in a way which results in an order of implementation which is less indicative of the relative quality of projects.

Upon financial integration, financial distortions affect global volatility patterns through two related channels: first, financial distortions determine the sensitivity of output to liquidity supply. Output in emerging markets is more sensitive to liquidity supply, because the projects implemented at the margin have high returns compared to the projects implemented infra-marginally. Second, in the integrated economy, the higher level of distortion in emerging markets causes them to absorb a larger fraction of the volatility of global liquidity supply. As a result, financial integration increases liquidity-driven output volatility in emerging markets, and decreases liquidity-driven output volatility in developed markets.

In the long run, a global environment in which liquidity is cheap is conducive to a deterioration in the financial system in the developed world. In the integrated economy, a deterioration in the quality of the financial system has a mixed effect on output in developed economies: in the absence of large liquidity or TFP shocks, it serves to increase output and to reduce output volatility. However, the response of output to large adverse liquidity shocks is amplified. This offers a modified view regarding the long run effect of financial integration on developed market output volatility. As the quality of financial institutions adjusts downward, the economy

becomes less sensitive to normal fluctuations, but more adversely affected by large shocks.

# Bibliography

Daron Acemoglu and Fabrizio Zilibotti. Was Prometheus unbound by chance? risk, diversification, and growth. *Journal of Political Economy*, 105(4):709–751, August 1997.

Susan Athey. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, February 2002.

Robert J. Barro. Rare disasters and asset markets in the twentieth century. *The Quarterly Journal of Economics*, 121(3):823–866, August 2006.

Fernando A. Broner and Jaume Ventura. Rethinking the effects of financial liberalization. unpublished manuscript, 2010.

Fernando A. Broner, Guido Lorenzoni, and Sergio L. Schmukler. Why do emerging economies borrow short term? *Journal of the European Economic Association*, Forthcoming. Forthcoming.

Markus K. Brunnermeier. Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives*, 23(1):77–100, Winter 2009.

Jeremy I. Bulow, John D. Geanakoplos, and Paul D. Klemperer. Multimarket oligopoly: Strategic substitutes and complements. *The Journal of Political Economy*, 93(3):488–511, June 1985.

Ricardo J. Caballero and Arvind Krishnamurthy. International and domestic collateral constraints in a model of emerging market crises. *Journal of Monetary Economics*, 48:513–548, 2001.

- Ricardo J. Caballero and Stavros Panageas. A quantitative model of sudden stops and external liquidity management. NBER Working Paper No. W11293, 2005.
- Ricardo J. Caballero, Kevin Cowan, and Jonathan Kearns. Fear of sudden stops: Lessons from australia and chile. *The Journal of Policy Reform*, 8(4):313–354, December 2005.
- Ricardo J. Caballero, Emmanuel Farhi, and Pierre-Olivier Gourinchas. An equilibrium model of ‘global imbalances’ and low interest rates. *American Economic Review*, 98(1):358–393, 2008.
- Roberto Chang and Andres Fernandez. On the sources of aggregate fluctuations in emerging economies. unpublished manuscript, 2010.
- Arnaud Costinot. An elementary theory of comparative advantage. *Econometrica*, 77(4):1165–1192, July 2009.
- Asli Demirguc-Kunt and Enrica Detragiache. Financial liberalization and financial fragility. World Bank Policy Research Working Paper No. 1917, 1999.
- Benjamin Eden. Does the bailout of banks imply higher future taxes: the fiscal implication of replacing one bubble asset with another. unpublished manuscript, 2009.
- Emmanuel Farhi and Jean Tirole. Bubbly liquidity. unpublished manuscript, 2010.
- Jesus Fernandez-Villaverde, Pablo Guerron-Quintana, Juan F. Rubio-Ramirez, and Martin Uribe. Risk matters: The real effects of volatility shocks. unpublished manuscript, 2009.
- Alessandra Fogli and Fabrizio Perri. The ‘great moderation’ and the us external imbalance. NBER Working Paper No. W12708, 2006.
- Ana Fostel and John Geanakoplos. Leverage cycles and the anxious economy. *American Economic Review*, 98(4):1211–1244, September 2008.

- Jordi Gali and Luca Gambetti. On the sources of the great moderation. *American Economic Journal: Macroeconomics*, 1(1):26–57, 2009.
- Mark Gertler and Nobuhiro Kiyotaki. Financial intermediation and credit policy in business cycle analysis. In Benjamin Friedman and Michael Woodford, editors, *Handbook of Monetary Economics*. Elsevier, 2010. forthcoming.
- Gary Gorton. The panic of 2007. In *Maintaining Stability in a Changing Financial System*. Symposium sponsored by The Federal Reserve of Kansas City at Jackson Hole, Wyoming, 2008.
- Jeremy Greenwood and Boyan Jovanovic. Financial development, growth, and the distribution of income. *Journal of Political Economy*, 98(5):1076–1107, October 1990.
- Dalia Hakura. Output volatility and large output drops in emerging market and developing countries. IMF Working Paper No. 07/114, 2007.
- Robert E. Hall. The high sensitivity of economic activity to financial frictions. unpublished manuscript, 2009.
- Bengt Holmstrom. Comment on: ‘the panic of 2007’, by gary gorton. In *Maintaining Stability in a Changing Financial System*. Symposium sponsored by The Federal Reserve of Kansas City at Jackson Hole, Wyoming, 2008.
- Ian Jewitt. Risk aversion and the choice between risky prospects: The preservation of comparative statics results. *The Review of Economic Studies*, 54(1):73–85, January 1987.
- Samuel Karlin and Yosef Rinott. Classes of orderings of measures and related correlation inequalities i. multivariate totally positive distributions. *Journal of Multivariate Analysis*, 10:467–498, 1980.
- Benjamin J. Keys, Tanmoy K. Mukherjee, Amit Seru, and Vikrant Vig. Did securitization lead to lax screening? evidence from subprime loans. EFA 2008 Athens Meetings Paper. Available at SSRN: <http://ssrn.com/abstract=1093137>, 2008.

- Nobuhiro Kiyotaki and John Moore. Credit cycles. *Journal of Political Economy*, 105(2):211–248, 1997.
- Miklos Koren and Silvana Tenreyro. Technological diversification. unpublished manuscript, 2009.
- Alberto Martin and Jaume Ventura. Theoretical notes on bubbles and the current crisis. unpublished manuscript, 2010.
- Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, September 1982.
- Pablo A. Neumeyer and Fabrizio Perri. Business cycles in emerging economies: the role of interest rates. *Journal of Monetary Economics*, 52(2):345–380, 2005.
- Maurice Obstfeld. Risk-taking, global diversification, and growth. *American Economic Review*, 84(5):1310–1329, December 1994.
- Raghuram G. Rajan. Why bank credit policies fluctuate: A theory and some evidence. *The Quarterly Journal of Economics*, 109(2):399–441, May 1994.
- Robert Shimer and Lones Smith. Assortative matching and search. *Econometrica*, 68(2):343–369, March 2000.
- Joseph E. Stiglitz. Capital market liberalization, economic growth, and instability. *World Development*, 28(6):1075–1086, 2000.
- Jean Tirole. Asset bubbles and overlapping generations. *Econometrica*, 53(6):1499–1528, November 1985.
- Martin Uribe and Vivian Z. Yue. Country spreads and emerging countries: Who drives whom? *Journal of International Economics*, 69:6–36, 2006.



## 1.8 Appendix A: Figures

The diverging trends of output volatility in emerging and developed markets. The series are taken from the World Economic Outlook, April, 2009. The series labeled “Emerging markets” corresponds to the series “Emerging and developing economies”, and the series labeled “Developed economies” corresponds to the series “Advanced economies”.

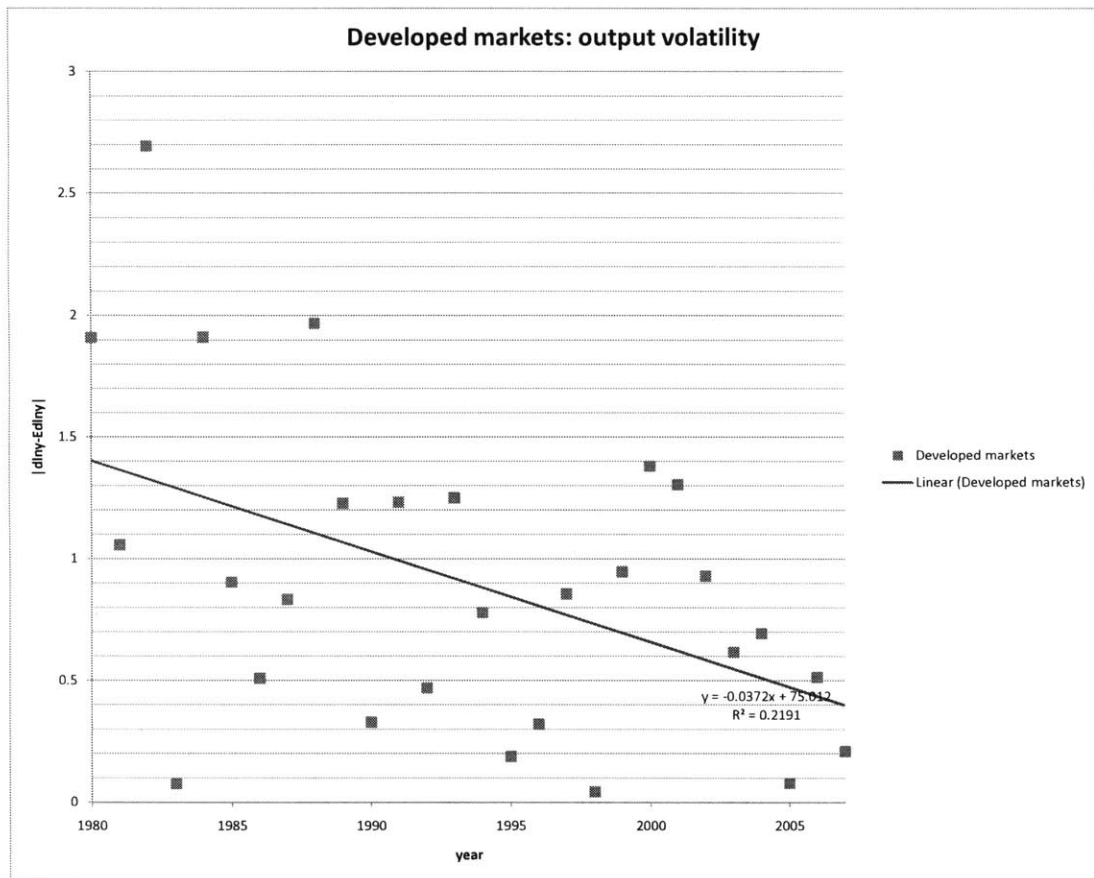


Figure 1-4: Output volatility trend in developed economies

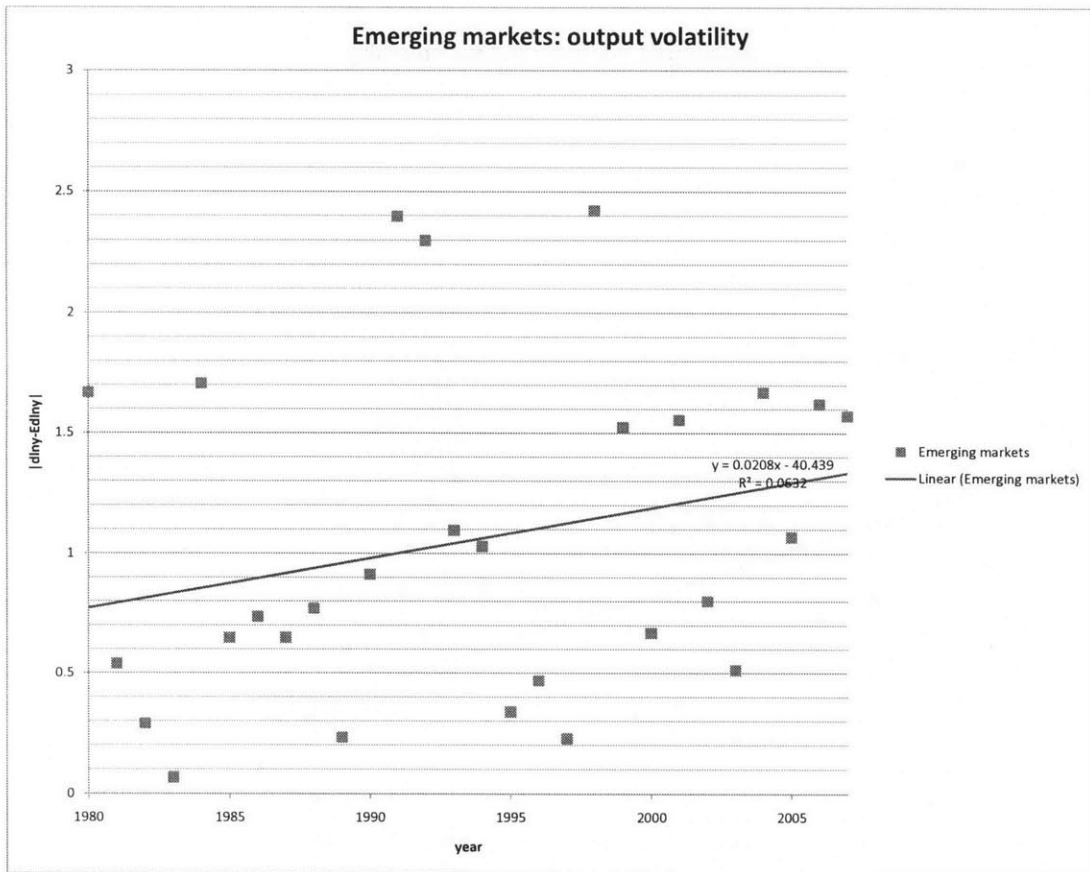


Figure 1-5: Output volatility trend in emerging economies

## 1.9 Appendix B: Proofs

### 1.9.1 Proof of Lemma 1

Let  $\phi > \phi'$  and consider the function:

$$Y(N, \phi') - Y(N, \phi) \quad (1.35)$$

Recall that this function is 0 for  $N = 0$  and for  $N = 1$ . Taking a first order condition yields:

$$y(N, \phi') - y(N, \phi) = 0 \quad (1.36)$$

This point is a local maximum, as the second derivative is negative by Assumption 1. To see this, using equation 1.36:

$$\frac{\frac{\partial y(N, \phi')}{\partial N} - \frac{\partial y(N, \phi)}{\partial N}}{y(N, \phi)} = \frac{\frac{\partial y(N, \phi')}{\partial N}}{y(N, \phi')} - \frac{\frac{\partial y(N, \phi)}{\partial N}}{y(N, \phi)} = \frac{\partial \ln y(N, \phi')}{\partial N} - \frac{\partial \ln y(N, \phi)}{\partial N} < 0 \quad (1.37)$$

The inequality stems from Assumption 1.

Thus, the function reaches a maximum at the point  $y(N, \phi') = y(N, \phi)$ . The minima of this function are therefore at  $N = 0$  and  $N = 1$ , at which the function takes the value of 0. It follows that the function is always weakly greater than 0:

$$Y(N, \phi') - Y(N, \phi) \geq 0 \quad (1.38)$$

Which concludes the proof.

### 1.9.2 Proof of Lemma 2

Begin by considering the case  $\phi \leq 1$ . First, note that for every  $N$ , there is a threshold  $\tilde{N}$  such that project  $(x, b)$  is implemented if and only if  $x \leq \tilde{N}$  and  $b \leq \tilde{N}$ . For  $\tilde{N} < \phi$ ,

$$N = Pr(x \leq \tilde{N}, b \leq \tilde{N}) = \tilde{N} \cdot \frac{1}{\phi} \tilde{N} = \frac{1}{\phi} \tilde{N}^2 \Rightarrow \tilde{N} = \sqrt{\phi} \sqrt{N} \quad (1.39)$$

Note that  $\tilde{N} < \phi$  if and only if  $N < \phi$ . For  $N > \phi$ , it is easy to see that  $\tilde{N} = N$ .

For  $N < \phi$ , output is given by the following expression:

$$Y(\tilde{N}) = \int_0^{\tilde{N}} Pr(b \leq \tilde{N}) Ag(x) dx = \frac{1}{\phi} \tilde{N} \int_0^{\tilde{N}} Ag(x) dx = \frac{A}{\phi} \tilde{N}^{1+\alpha} \quad (1.40)$$

It follows that, for  $N < \phi$ :

$$Y(N, \phi) = A\phi^{\frac{\alpha-1}{2}} N^{\frac{1+\alpha}{2}} \quad (1.41)$$

The derivative of  $Y$  with respect to  $N$  is therefore given by:

$$y(N, \phi) = \frac{1+\alpha}{2} A\phi^{\frac{\alpha-1}{2}} N^{\frac{\alpha-1}{2}} \quad (1.42)$$

$$\Rightarrow \ln y(N, \phi) = \ln\left(\frac{1+\alpha}{2} A\right) - \frac{1-\alpha}{2} \ln \phi - \frac{1-\alpha}{2} \ln N \quad (1.43)$$

The derivative of above with respect to  $N$  is:

$$\frac{\partial \ln y(N, \phi)}{\partial N} = -\frac{1-\alpha}{2N} \quad (1.44)$$

The above does not depend on  $\phi$  as long as  $N < \phi$ ; the log supermodularity condition is trivially satisfied. However, Note that for a higher  $\phi$ , there are more values of  $N$  such that  $N < \phi$ . Let there be  $\phi$  and  $\phi'$  such that  $\phi' < N < \phi$ . For  $\phi'$ , it is easy to see that the derivative of  $\ln y(N, \phi')$  is given by the following expression:

$$y(N, \phi') = A\alpha N^{-(1-\alpha)} \Rightarrow \ln y(N, \phi') = \ln(A\alpha) - (1-\alpha) \ln N \quad (1.45)$$

$$\Rightarrow \frac{\partial \ln y(N, \phi')}{\partial N} = -\frac{1-\alpha}{N} < -\frac{1-\alpha}{2N} = \frac{\partial \ln y(N, \phi)}{\partial N} \quad (1.46)$$

Thus, the derivative of  $\ln y$  with respect to  $N$  is higher in the more distorted economy  $\phi$ , in accordance with Assumption 1.

Consider now the range  $\phi \geq 1$ . In this range, for  $\tilde{N} < 1$ ,  $N$  is given by equation 1.39, output is given by equation 1.40 and  $\frac{\partial \ln y}{\partial N}$  is given by equation 1.44. In this

range,  $\frac{\partial \ln y}{\partial N}$  is constant with respect to  $\phi$ , so the log supermodular condition is trivially satisfied.

Note that  $\tilde{N} < 1$  if and only if  $\sqrt{\phi N} < 1$ , or  $N < \frac{1}{\phi}$ . This condition is violated for more values of  $N$  if  $\phi$  is larger. For  $N > \frac{1}{\phi}$ , the marginal product of working capital is constant; the collateral constraint is binding for all implemented projects, so the productivity of the projects implemented with each unit of working capital is the same. Thus, in the range  $N > \frac{1}{\phi}$ ,

$$\frac{\partial \ln y(N, \phi)}{\partial N} = 0 > -\frac{1 - \alpha}{2N} \quad (1.47)$$

The right hand side is equal to the derivative of  $\ln y$  for the case  $N < \frac{1}{\phi}$ . It follows that the log supermodularity condition is satisfied for  $\frac{1}{\phi} < N < \frac{1}{\phi'}$ .

### 1.9.3 Proof of Lemma 3

Output is given by the following expression:

$$Y(N, \phi) = \sum_{i=0}^{\phi} \int_{\frac{i}{\phi}}^{\frac{i}{\phi} + \frac{N}{\phi}} Ag(x) dx \quad (1.48)$$

This is because each local bank has  $\frac{N}{\phi}$  units of liquidity to allocate, and uses it to implement the first  $\frac{N}{\phi}$  in the sample of projects available to it.

The marginal product of working capital is given by:

$$y(N, \phi) = \frac{1}{\phi} \sum_{i=0}^{\phi} Ag\left(\frac{i}{\phi} + \frac{N}{\phi}\right) = \frac{A}{\phi} \sum_{i=0}^{\phi} e^{-\left(\frac{i}{\phi} + \frac{N}{\phi}\right)} \quad (1.49)$$

$$= \frac{Ae^{-\frac{N}{\phi}}}{\phi} \sum_{i=0}^{\phi} e^{-\frac{i}{\phi}} \quad (1.50)$$

It follows that:

$$\ln y(N, \phi) = \ln\left(\frac{A}{\phi} \sum_{i=0}^{\phi} e^{-\frac{i}{\phi}}\right) + \ln e^{-\frac{N}{\phi}} = c - \frac{N}{\phi} \quad (1.51)$$

The derivative of above with respect to  $N$  is  $-\frac{1}{\phi}$ , which is increasing in  $\phi$ . The log

supermodularity condition is satisfied, in accordance with Assumption 1.

### 1.9.4 Proof of Proposition 1

To show that  $Y(N, \phi)$  is log supermodular, write  $Y(N, \phi)$  as:

$$Y(N, \phi) = \int_0^\infty 1_{[0, N]}(n) y(n, \phi) dn \quad (1.52)$$

Where  $1_{[0, N]}$  denotes the indicator function which takes a value 1 over the interval  $[0, N]$  (and 0 elsewhere).

Recall the definition of log supermodularity as it appears in Costinot [2009], which allows for 0 values:

**Definition 1** For  $X \subset \mathbb{R}^w$ , a function  $h : X \rightarrow \mathbb{R}^+$  is log supermodular if for all  $z, z' \in X$ ,

$$h(\max(z_1, z'_1), \dots, \max(z_m, z'_m)) h(\min(z_1, z'_1), \dots, \min(z_m, z'_m)) \geq h(z) h(z') \quad (1.53)$$

**Claim 1** The function  $h(N, n, \phi) = 1_{[0, N]}(n)$  is log supermodular in  $N, n$ , and  $\phi$ .

To see this, note that both sides of the inequality in 1.53 can be either 0 or 1, and consider the case in which the left hand side of the inequality is 0:

$$h(\max(N, N'), \max(n, n'), \max(\phi, \phi')) h(\min(N, N'), \min(n, n'), \min(\phi, \phi')) = 0 \quad (1.54)$$

$$1_{[0, \max(N, N')]}(\max(n, n')) 1_{[0, \min(N, N')]}(\min(n, n')) = 0 \quad (1.55)$$

Assume without loss of generality that  $\max(n, n') = n$ . From the above equality,  $1_{[0, \max(N, N')]}(\max(n, n')) = 0$  or  $1_{[0, \min(N, N')]}(\min(n, n')) = 0$ . Assume  $1_{[0, \max(N, N')]}(\max(n, n')) = 0$ . Thus,

$$n > \max(N, N') \Rightarrow n > N \Rightarrow 1_{[0, N]}(n) = 0 \Rightarrow 1_{[0, N]}(n) 1_{[0, N']}(n') = 0 \quad (1.56)$$

Assume instead that  $1_{[0, \min(N, N')]}(\min(x, x')) = 0$ . There are two cases: if  $\min(N, N') = N'$ ,

$$n' > N' \Rightarrow 1_{[0, N']}(n') = 0 \Rightarrow 1_{[0, N]}(n)1_{[0, N']}(n') = 0 \quad (1.57)$$

if, instead,  $\min(N, N') = N$ , since  $\max(n, n') = n$ ,

$$n' > N \Rightarrow n > N \Rightarrow 1_{[0, N]}(n) = 0 \Rightarrow 1_{[0, N]}(n)1_{[0, N']}(n') = 0 \quad (1.58)$$

Thus, log supermodularity is satisfied.

The assumption that  $y(n, \phi)$  is log supermodular in  $(n, \phi)$  implies trivially that it is log supermodular as a function of  $(N, n, \phi)$ .

Since the product of two log supermodular functions is log supermodular, and the integral of a log supermodular function is log supermodular<sup>15</sup>, it follows that  $Y(N, \phi)$  is log supermodular. Thus, by log supermodularity,  $\frac{\partial \ln Y(N, \phi)}{\partial N}$  is increasing in  $\phi$ .

The second part of the proposition builds on the first part:

$$\frac{Y(N, \phi)}{N} = Y(N, \phi)N^{-1} \quad (1.59)$$

Since  $Y(N, \phi)$  is log supermodular, and  $N^{-1}$  is trivially log supermodular, it follows that average productivity is log supermodular as a product of two log supermodular functions. Thus, by log supermodularity,  $\frac{\partial \ln(\frac{Y(N, \phi)}{N})}{\partial N}$  is increasing in  $\phi$ . Since the derivative is negative, this implies that the sensitivity of average productivity to the level of working capital is higher in less distorted economies.

### 1.9.5 Proof of Propositions 2 and 3

1. To show that financial integration exacerbates volatility differences. I begin by showing that (at autarkic working capital levels) working capital is more sensitive to the price of liquidity in emerging markets. Since output is more sensitive to working capital (Proposition 1), and since, by assumption, mean

---

<sup>15</sup>For proof see Karlin and Rinott [1980].

levels of working capital remain unchanged following financial integration, the proposition follows. Since it is assumed that  $f$  is decreasing, the working capital level is such that the marginal product of working capital is equal to its market price  $r$ :

$$y(N, \phi) = r \quad (1.60)$$

It follows that the derivative of  $N$  with respect to  $r$  is:

$$\frac{\partial N}{\partial r} = \frac{1}{\frac{\partial y}{\partial N}} = \frac{1}{\frac{\partial y(N, \phi)}{\partial N}} \quad (1.61)$$

Denote:

$$f(N, \phi) = \int_0^1 \sigma_\phi(N, x) g(x) dx \quad (1.62)$$

Note that  $y(N, \phi; A) = Af(N, \phi)$ . The assumption that  $y$  is log supermodular implies trivially that  $f$  is log supermodular.

Log supermodularity implies that:

$$\frac{A_{em} f_1(N, \phi_{em})}{A_{em} f(N, \phi_{em})} = \frac{f_1(N, \phi_{em})}{f(N, \phi_{em})} > \frac{f_1(N, \phi_d)}{f(N, \phi_d)} = \frac{A_d f_1(N, \phi_d)}{A_d f(N, \phi_d)} \quad (1.63)$$

Since, by assumption,  $A_{em} f(N, \phi_{em}) = A_d f(N, \phi_d)$ , it follows that

$$A_{em} f_1(N, \phi_{em}) > A_d f_1(N, \phi_d) \quad (1.64)$$

It follows that:

$$\frac{\partial N_d}{\partial r} = \frac{1}{A_d f_1(N, \phi_d)} > \frac{1}{A_{em} f_1(N, \phi_{em})} = \frac{\partial N_{em}}{\partial r} \quad (1.65)$$

Since the response of  $N$  to  $r$  is negative, it follows that  $N_{em}$  is more sensitive to changes in  $r$ . By construction,  $N_{em}$  and  $N_d$  are similarly affected by shocks to relative productivity (which cause a substitution between  $N_d$  and  $N_{em}$ ). Thus, since these are the only two sources of variation in  $N$ , under the assumption



that  $Q_i = N_i^a$  are identically distributed,

$$\text{var}(\ln N_{em}) - \text{var}(\ln N_d) > \text{var}(\ln N_{em}^a) - \text{var}(\ln N_d^a) = 0 \quad (1.66)$$

To see that financial integration exacerbates the difference in output volatility, note that:

$$\text{var}(\ln Y) = \frac{\partial \ln Y}{\partial \ln N} \text{var}(\ln N) + \text{var}(\ln A) \quad (1.67)$$

Decompose volatility in  $N$  into volatility conditional on  $Q$  shocks and volatility conditional on relative TFP shocks:

$$\text{var}(\ln N) = \text{var}(\ln N|Q) + \text{var}(\ln N|A) \quad (1.68)$$

$$\text{var}(\ln Y_{em}) - \text{var}(\ln Y_d) - (\text{var}(\ln A_{em}) - \text{var}(\ln A_d)) = \quad (1.69)$$

$$\frac{\partial \ln Y_{em}}{\partial \ln N} (\text{var}(\ln N_{em}|Q) + \text{var}(\ln N_{em}|A)) - \frac{\partial \ln Y_d}{\partial \ln N} (\text{var}(\ln N_d|Q) + \text{var}(\ln N_d|A)) \quad (1.70)$$

$$= \frac{\partial \ln Y_{em}}{\partial \ln N} \text{var}(\ln N_{em}|A) - \frac{\partial \ln Y_d}{\partial \ln N} \text{var}(\ln N_d|A) + \quad (1.71)$$

$$\text{var}(\ln N|Q) \left( \frac{\partial \ln Y_{em}}{\partial \ln N} - \frac{\partial \ln Y_d}{\partial \ln N} \right) \quad (1.72)$$

Since the last term is positive, the above is greater than the first term in the above expression:

$$> \frac{\partial \ln Y_{em}}{\partial \ln N} \text{var}(\ln N_{em}|A) - \frac{\partial \ln Y_d}{\partial \ln N} \text{var}(\ln N_d|A) \quad (1.73)$$

The following lemma will be useful to conclude the proof:

**Lemma 5** (a)  $\text{var}(\ln N_{em}|A) > \text{var}(\ln N_{em}^a)$

(b)  $\text{var}(\ln N_d|A) < \text{var}(\ln N_d^a)$

**Proof:** To see this, consider a benchmark in which economy  $i$  integrates with an economy with an **identical** sensitivity to  $r$  shocks. In this hypothetical case, shocks to domestic and foreign liquidity supply adjust equally between the two countries. Using the assumption that  $Q_i$  are perfectly correlated, it follows that:

$$\text{var}(N_i) = \text{var}\left(\frac{1}{2}(Q_i + Q_j)\right) = \frac{1}{4}(\text{var}(2Q_i)) = \text{var}(Q_i) = \text{var}(N_i^a) \quad (1.74)$$

The lemma above is proved using comparative statics with this benchmark. Since shocks to liquidity supply have a greater effect on emerging markets, the volatility of  $\text{var}(N_{em})$  is greater than this benchmark, and since shocks to liquidity supply have a smaller effect on developed markets, the volatility of  $\text{var}(N_d)$  is lower than this benchmark. As mean liquidity levels stay the same, the lemma (as stated in logs) immediately follows.

Using the above lemma, the expression in equation 1.73 is greater than the expression below, in which  $\text{var}(\ln N_i|A)$  are replaced with autarkic values:

$$(1.73) > \frac{\partial \ln Y_{em}}{\partial \ln N} \text{var}(\ln N_{em}^a) - \frac{\partial \ln Y_d}{\partial \ln N} \text{var}(\ln N_d^a) \quad (1.75)$$

$$= \text{var}(\ln Y_{em}^a) - \text{var}(\ln Y_d^a) - (\text{var}(\ln A_{em}) - \text{var}(\ln A_d)) \quad (1.76)$$

$$\Rightarrow \text{var}(\ln Y_{em}) - \text{var}(\ln Y_d) > \text{var}(\ln Y_{em}^a) - \text{var}(\ln Y_d^a) \quad (1.77)$$

Which concludes the proof.

2. (a) To see that financial integration increases volatility of working capital in the emerging market region, note that both  $\text{var}(N_{em}|A) > \text{var}(N_{em}^a|A) = 0$ , and  $\text{var}(N_{em}|Q) > \text{var}(N_{em}^a|Q)$ . To see that financial integration increases the volatility of output in the emerging market region, note that

since  $var(A)$  remains the same, and since the level  $N$  is unchanged, from equation 1.67 output volatility increases as well.

- (b) The effect of financial integration on  $var(N_d)$  is ambiguous: the volatility of  $N_d$  conditional on holding technology levels constant is smaller, so  $var(N_d|A) < var(N_d^a|A)$ . However, through the standard RBC channel,  $var(N_d|Q) > var(N_d^a|Q) = 0$ . If shocks to relative TFP are sufficiently small, the first effect dominates so  $var(N_d) < var(N_d^a)$ . In this case, from equation 1.67 (and the assumption that the mean level of  $N$  is unchanged), it follows that output volatility decreases as well.

## 1.9.6 Proof of Proposition 4

**Lemma 6** *Consider the problem:*

$$\max_{\phi} \int_0^N y(N', \phi) dN' - \lambda(\phi) \quad (1.78)$$

*The solution  $\phi^*$  is increasing in  $N$  for some range  $N \in (\tilde{N}, 1]$ .*

Using the proof of Lemma 1, for each pair  $\phi_i > \phi_j$  there is a point  $0 < N_{i,j} < 1$  such that  $Y(N, \phi)$  has increasing differences for  $N > N_{i,j}$  and  $\phi \in \{\phi_i, \phi_j\}$ . Let  $\tilde{N}$  denote the maximum of these  $N_{i,j}$ . In the region  $(\tilde{N}, 1]$ , there are increasing differences for all  $\phi$  within the set of possible values. By Sundaram (1996),  $\phi^*$  is increasing in  $N$  on that region.

Note that, for any  $r$ ,  $\phi^*$  is the solution to:

$$\max_{\phi} \int_0^N y(N', \phi) dN' - \lambda(\phi) - rN \quad (1.79)$$

This is because  $rN$  is a constant in this problem.

Now, assume that  $N \in (\tilde{N}, 1]$ . In this range, a drop in  $r$  is associated with an increase in  $N^*$ . It follows that  $\phi^*$  increases.

### 1.9.7 Proof of Proposition 5

For what follows, denote by  $\phi_d^w$  the level of distortions in the deteriorated “weak” financial system, and by  $\phi_d^i$  the autarkic “intact” level of financial distortions. Normalize  $\lambda(\phi_d^w) = 0$  and  $\lambda(\phi_d^i) = \lambda$ .

I begin with the first part of the proposition. By equations 1.36 and 1.37, there is a unique point  $N_0 \in (0, 1)$  such that:

$$y_d(N_0, \phi_d^i) = y_d(N_0, \phi_d^w) \quad (1.80)$$

For every  $N < N_0$ ,  $y(N, \phi_d^i) > y(N, \phi_d^w)$  and for every  $N > N_0$ ,  $y(N, \phi_d^i) < y(N, \phi_d^w)$ .

Recall that in this setup, prior to financial integration the financial system in the developed world was endogenously intact, and it endogenously deteriorated following a drop in  $r$ . It turns out that these dynamics are possible only if  $r$  is such that working capital levels exceed  $N_0$  (when the financial system is intact and the economy is integrated):

**Lemma 7** *There is an endogenous weakening of the financial system (in the developed world) only if  $r \leq y(N_0, \phi_d^i) = y(N_0, \phi_d^w)$ .*

Here,  $r$  denotes the rate of return on working capital when the financial system in the developed world is intact.

**Proof:** Assume  $r > y(N_0, \phi_d^i)$ , and it will be shown that in this case there is no endogenous financial deterioration. Denote by  $r^a$  the return to working capital under autarky in the developed world. We know that given  $r^a$ , the optimal choice of financial quality is  $\phi_d = \phi_d^i$ :

$$\pi(r^a, \phi_d^i) > \pi(r^a, \phi_d^w) \quad (1.81)$$

$$\Rightarrow \max_N \int_0^N y(N', \phi_d^i) dN' - r^a N - \lambda > \max_N \int_0^N y(N', \phi_d^w) dN' - r^a N \quad (1.82)$$

The standard optimality condition is  $y(N, \phi_d) = r^a$ . Denote this  $N$  by  $N^*(r, \phi_d)$ .

Thus, the inequality in equation 1.82 can be rewritten as:

$$\int_{r^a}^{\infty} (y(N^*(r', \phi_d^i), \phi_d^i) - y(N^*(r', \phi_d^w), \phi_d^w)) dr' > \lambda \quad (1.83)$$

From the assumption that  $r > y(N_0, \phi_d^i)$ , it follows that  $r^a > y(N_0, \phi_d)$ , because  $r^a > r$  (there is a drop in  $r$  upon financial integration). For  $r > y(N_0, \phi_d^i)$ , it is also the case that:

$$N^*(r, \phi_d^i) > N^*(r, \phi_d^w) \quad (1.84)$$

This is because, since  $N^*(r, \phi_d^i) < N_0$ ,

$$y(N^*(r, \phi_d^i), \phi_d^w) < y(N^*(r, \phi_d^i), \phi_d^i) \quad (1.85)$$

Thus, since  $y$  is decreasing in  $N$ , the inequality in equation 1.84 holds true.

Note that equation 1.85 holds for any  $y(N_0, \phi_d) < r < r^a$ . Thus,

$$\int_0^{N^*(r, \phi_d^i)} y(N', \phi_d^i) dN' - rN^*(r, \phi_d^i) - \int_0^{N^*(r, \phi_d^w)} y(N', \phi_d^w) dN' - rN^*(r, \phi_d^w) = \quad (1.86)$$

$$\int_r^{\infty} y(N^*(r', \phi_d^i), \phi_d^i) dr' - \int_r^{\infty} y(N^*(r', \phi_d^w), \phi_d^w) dr' = \quad (1.87)$$

$$\int_{r^a}^{\infty} (y(N^*(r', \phi_d^i), \phi_d^i) - y(N^*(r', \phi_d^w), \phi_d^w)) dr' + \quad (1.88)$$

$$\int_r^{r^a} (y(N^*(r', \phi_d^i), \phi_d^i) - y(N^*(r', \phi_d^w), \phi_d^w)) dr' \quad (1.89)$$

By equation 1.83, the first term is greater than  $\lambda$ . The second term is positive by equation 1.85. Thus, the sum above is greater than  $\lambda$ . It follows that choosing  $\phi_d = \phi_d^i$  is still preferable to choosing  $\phi_d = \phi_d^w$ . I conclude that an endogenous financial deterioration is not possible if  $r > y(N_0, \phi_d^i)$ .

To conclude the proof, note that for  $r < y(N_0, \phi_d^i)$ ,

$$N^*(r, \phi_d^i) < N^*(r, \phi_d^w) \quad (1.90)$$

Thus, the deterioration in the quality of the financial system is associated with an increase in  $N$ .

To see the second part of the proposition, note that for  $\frac{A_d}{A_{em}}$  sufficiently large,

$$y_d(1, \phi_d^w) > y_{em}(1, \phi_{em}) \quad (1.91)$$

From continuity, there exists  $\epsilon > 0$  such that for every  $1 - \epsilon < N < 1$ ,

$$y_d(1, \phi_d^w) > y_{em}(N, \phi_{em}) \quad (1.92)$$

It follows that for  $Q_w > 2 - \epsilon$ , the developed market will be satiated with working capital. Output is weakly higher than in the intact case (note that the conditions under which the weakened financial system is satiated with working capital are weaker than the conditions under which the intact economy is satiated with working capital, because, since  $N_0 < 1$ ,  $y_d(1, \phi_d^w) > y_d(1, \phi_d^i)$ ). Under satiation, the developed economy is stable with respect to small shocks to the global supply of working capital, as the economy remains satiated as long as  $Q_w$  satisfies:

$$y_d(1, \phi_d^w) > y_{em}(Q_w - 1, \phi_{em}) \quad (1.93)$$

Similarly, the level of working capital is unaffected by shocks to relative TFP. The volatility of output is therefore smaller both with respect to liquidity supply shocks and with respect to shocks to relative TFP.

### 1.9.8 Proof of Lemma 4

Recall the notation from the proof of Proposition 5:  $\phi_d^w$  denotes the level of distortions in the deteriorated “weak” financial system, and  $\phi_d^i$  denotes the autarkic “intact” level of financial distortions.

Given  $\phi_d$ , world output maximization solves:

$$Y_w(\phi_d) = \max_{N_d, N_{em}} Y_w(\phi_d, N_d, N_{em}) = \quad (1.94)$$

$$\max_{N_d, N_{em}} A_d \int_0^{N_d} y(N', \phi_d) dN' + A_{em} \int_0^{N_{em}} y(N', \phi_{em}) dN' \quad (1.95)$$

s.t.

$$N_d + N_{em} = Q_w \quad (1.96)$$

$$N_i \leq 1 \quad (1.97)$$

Under autarky there is an additional constraint which is:

$$N_i = Q_i \quad (1.98)$$

It is easy to see that for any *given* couple  $(N_d, N_{em})$  which satisfy constraints 1.96-1.97 or 1.96-1.98, the output produced is higher when the financial system in the developed market is intact:

$$Y_w(\phi_d^i, N_d, N_{em}) \geq Y_w(\phi_d^w, N_d, N_{em}) \quad (1.99)$$

Denote by  $(N_d^*(\phi_d), N_{em}^*(\phi_d))$  the optimal allocation of working capital given  $\phi_d$ . From the optimality of  $N_i^*(\phi_d^i)$ :

$$Y_w(\phi_d^i) = Y_w(\phi_d^i, N_d^*(\phi_d^i), N_{em}^*(\phi_d^i)) \geq Y_w(\phi_d^i, N_d^*(\phi_d^w), N_{em}^*(\phi_d^w)) \quad (1.100)$$

And, from equation 1.99:

$$Y_w(\phi_d^i, N_d^*(\phi_d^w), N_{em}^*(\phi_d^w)) \geq Y_w(\phi_d^w, N_d^*(\phi_d^w), N_{em}^*(\phi_d^w)) = Y_w(\phi_d^w) \quad (1.101)$$

It follows that world output is higher when the financial system in the developed market is intact:

$$\Rightarrow Y_w(\phi_d^i) \geq Y_w(\phi_d^w) \quad (1.102)$$

### 1.9.9 Proof of Proposition 6

Recall the notation from the proof of Proposition 5:  $\phi_d^w$  denotes the level of distortions in the deteriorated “weak” financial system, and  $\phi_d^i$  denotes the autarkic “intact” level of financial distortions.

Let  $N_0$  be given by the condition in equation 1.36:

$$y(N_0, \phi_d^i) = y(N_0, \phi_d^w) \quad (1.103)$$

Consider a shock to the global supply of liquidity such that  $Q_w < N_0$ . In this range, the marginal product of working capital in the developed world is higher under intact financial institutions. It follows that working capital is higher:

$$N_d^*(Q_w, \phi_d^i) > N_d^*(Q_w, \phi_d^w) \quad (1.104)$$

Since output decreases with the level of distortion by Lemma 1,

$$Y_d(N_d^*(Q_w, \phi_d^i), \phi_d^i) > Y_d(N_d^*(Q_w, \phi_d^i), \phi_d^w) \quad (1.105)$$

Since output is increasing in the level of working capital,

$$Y_d(N_d^*(Q_w, \phi_d^i), \phi_d^w) > Y_d(N_d^*(Q_w, \phi_d^w), \phi_d^w) \quad (1.106)$$

It follows that output is higher under intact financial institutions:

$$Y_d(N_d^*(Q_w, \phi_d^i), \phi_d^i) > Y_d(N_d^*(Q_w, \phi_d^w), \phi_d^w) \quad (1.107)$$

Similarly, a shock to relative TFP such that  $N_d^*(Q_w, \phi_d^i) < N_0$  is amplified by lack of high quality financial institutions.



## 1.10 Appendix C: A Monetary Model of Liquidity Supply

In this section, I develop a model of liquidity supply, in which liquidity supply fluctuations are caused by two primitive shocks: shocks to the money supply and shocks to consumers' risk aversion<sup>16</sup>. Recall that the model makes three assumptions about the distribution of liquidity supply:

1. The distribution of liquidity supply is unchanged by financial integration.
2. Liquidity supply is independent from both domestic and foreign productivity,  $A_i$ .
3. Liquidity supply is perfectly correlated across regions.

Consider a model in which labor (denoted  $L$ ) is the only productive input. Given  $L_i$  hired units of labor, output in country  $i$  is given by:

$$Y_i = A_i F_i(L_i) \tag{1.108}$$

In this model, liquidity is money used to hire labor; thus, in this formulation,  $F_i$  already captures the efficiency of the financial system in allocating liquidity (in other words,  $F$  is log supermodular in  $L$  and  $\phi$ ).

**Assumption 2**    1. *The price of a unit of labor in terms of money,  $w$ , is determined at the beginning of the period, before all shocks are realized. Agents agree to supply any amount of labor for the wage  $w$ .*

2. *The price of output is fixed within a period, and is normalized to one.*

---

<sup>16</sup>The analysis of Broner et al. [Forthcoming] suggests that shocks to the risk aversion of international investors is indeed an important driving source of supply driven volatility in emerging markets.

Agents live for one period and consume at the end of their lives. The preferences of agents in region  $i$  are given by:

$$E(u_i(c_i) - v_i(L_i)) \quad (1.109)$$

I assume that  $u_i(c_i)$  takes the following stark form:

$$u_i(c_i) = \begin{cases} c_i & \text{if } c \geq c_{0,i}; \\ -\infty & \text{otherwise.} \end{cases} \quad (1.110)$$

In this formulation,  $c_{0,i}$  captures the level of risk aversion of households in region  $i$ . To see this, consider the comparison between two agents, one denoted  $h$  with  $c_{0,h} = c_{0,h}$  and one denoted  $l$  with  $c_{0,l} < c_{0,h}$ .

**Definition 2** *Agent  $h$  is more risk averse than agent  $l$  if the following condition holds: for any certain consumption payment  $c$ , and any lottery  $q$ , if agent  $h$  prefers  $q$  over  $c$  then so does agent  $l$ .*

To see that, according to this standard definition, the condition  $c_{0,h} > c_{0,l}$  implies that agent  $h$  is more risk averse than agent  $l$ , note the following claim:

**Claim 2** *Agent  $i$  prefers a lottery  $q$  over a certainty payment of  $c$  if and only if the lottery  $q$  never delivers a payment of less than  $c_{0,i}$ , and  $E(q) > c$ .*

**Proof:** Trivially, if the above condition holds, the agent will prefer the lottery: if  $c > c_{0,i}$  he is risk neutral between the two lotteries, and if  $c \leq c_{0,i}$  his utility from consuming  $c$  is  $-\infty$  whereas it is positive given the lottery. If the above condition is violated, it means that one of the following holds: either  $E(q)$  delivers a payment of less than  $c_{0,i}$  with positive probability, or  $E(q) < c$ . If  $E(q)$  delivers a payment of less than  $c_{0,i}$  with positive probability, then the agent's expected utility from the lottery is  $-\infty$ , so it is not preferred over anything. If  $E(q) < c$ , but  $q$  always delivers a payment of more than  $c_{0,i}$ , then it follows that  $c > c_{0,i}$ , so the agent is risk neutral with respect to  $q$  and  $c$  and would prefer  $c$ .

Using this claim, it is easy to see that  $h$  is more risk averse than  $l$ , as the fact that  $q$  never delivers a payment of less than  $c_{0,h}$  implies that it never delivers a payment of less than  $c_{0,l} < c_{0,h}$ , so the set of lotteries and certainty payments in which  $h$  prefers the lottery is included in the set of lotteries and certainty payments in which  $l$  prefers the lottery.

Thus, we will think of  $c_0$  as the level of risk aversion of the agents.

**Corollary 2** *A higher level of  $c_{0,i}$  implies a higher level of risk aversion.*

### 1.10.1 Liquidity Supply in the Closed Economy

After wages are agreed upon and prices are set, the money supply,  $M_i$ , and the level of risk aversion,  $c_{0,i}$ , are realized. Agents can choose to hold their money in a safe ( $M_{h,i}$ ) or buy stocks in the productive sector ( $Q_i$ ):

$$M_i = Q_i + M_{h,i} \tag{1.111}$$

The level of  $Q_i$  is also the level of liquidity supply which can be used to hire workers. After the productivity shock  $A_i$  is realized, the financial system allocates the liquidity  $Q_i$  to domestic projects who use it to hire workers:

$$Q_i = w_i L_i \tag{1.112}$$

After production takes place but before workers are paid, there is a shock to the ability of the productive sector to make monetary transfers. With probability  $(1 - \theta)$ , this ability is *intact*; in this case, wages are paid and two rounds of consumption follow. In the first round, households use their money holdings (which include wage payments,  $w_i L_i$ , and money from the safe,  $M_{h,i}$ ) to buy output and consume. In the second round, monetary revenues from sales are redistributed to households as dividends (denoted  $d$  per share), and are used for consumption. The end of the period consumption is given by:

$$c_i = M_{h,i} + w_i L_i + dQ_i \tag{1.113}$$

With a small probability  $\theta$ , the ability of the productive sector to make monetary transfers *collapses*. This implies two things: first, wages are not paid to workers. Second, revenues from sales cannot be redistributed back to households, so no dividends are paid. There is only one round of consumption, in which households use the money which they had kept in the safe ( $M_{h,i}$ ) to buy consumption goods. It follows that the end of the period consumption is given by:

$$c_i = M_{h,i} \tag{1.114}$$

Simplistically, the shock to the ability of the productive sector to make monetary transfers can be thought of as a strike in the postal services: households who receive wage payments and dividend payments by mail (and are subject to a cash in advance constraint) do not receive these payments in time to consume before they die. Realistically, this shock is meant to capture a shock to the money supply, in which certain substitutes for money used by the productive sector are no longer valued<sup>17</sup>. For example, prior to the sub-prime crisis, mortgage backed securities were accepted by all as means of payment. In the sub-prime crisis, these securities turned into illiquid assets, and people were no longer willing to hold these assets without understanding the value of their components. Thus, it became harder to trade these assets for consumption goods, which made bonds and cash more desired<sup>18</sup>. In this model, agents hold bonds (“money in the safe”) precisely to insure against events of this kind<sup>19</sup>.

To summarize, the timing within a period is as follows:

1. The wage  $w_i$  and the price of output (normalized to 1) is set.
2. The initial money supply,  $M_{0,i}$ , and the level of risk aversion,  $c_{0,i}$ , are realized.

---

<sup>17</sup>See Eden [2009] for a complete development of this idea.

<sup>18</sup>See Holmstrom [2008] for a complete development of this idea.

<sup>19</sup>Results similar to those derived in this section can be derived in a more standard framework, in which agents hold money to insure against unemployment risk; however, this motive for hoarding liquidity seems less compelling as the mechanism through which shocks to risk aversion affect liquidity supply. Rather, in this formulation portfolio decisions are motivated by fear of large aggregate disasters, consistent with the view expressed in Barro [2006].

3. Agents use some of their money to buy stocks and effectively supply liquidity to the productive sector (and keep the rest in a safe).
4. The productivity shock,  $A_i$ , is realized.
5. The financial system allocates liquidity to projects, and production takes place.
6.
  - With prob.  $(1-\theta)$  (*intact*): wages are paid, followed by a first round of consumption in which agents trade their wage earnings ( $w_i L_i$ ) and their money holdings ( $M_{h,i}$ ) for consumption. The productive sector redistributes revenues from sales back to households in the form of dividends. A second round of consumption takes place in which the dividends are traded for consumption.
  - With prob.  $\theta$  (*collapse*): the productive sector loses its ability to make monetary transfers. Wages are not paid and revenues from sales cannot be redistributed back to households. There is therefore only one round of consumption, in which households use their money holdings ( $M_{h,i}$ ) to buy consumption goods.

## Equilibrium

**The portfolio decision.** Clearly, the agent will reserve at least enough money to finance  $c_{0,i}$  units of consumption in case of a collapse (as otherwise his expected utility is  $-\infty$ ):

$$M_{h,i} \geq c_{0,i} \tag{1.115}$$

Denote by  $d$  the realized dividend per share of the productive sector, and let  $Q_i^0$  be the equilibrium level of liquidity supply, which the individual agent takes as given.

Agents solve:

$$\max_{Q_i, M_{h,i}} c_i - v_i(L_i) \tag{1.116}$$

s.t.:

$$M_i = Q_i + M_{h,i} \tag{1.117}$$

$$c_i = \begin{cases} M_{h,i} + Q_i d + w_i L_i & \text{with prob. } (1 - \theta); \\ M_{h,i} & \text{with prob. } \theta. \end{cases} \quad (1.118)$$

$$M_{h,i} \geq c_{0,i} \quad (1.119)$$

Substituting in the first constraint, this problem can be rewritten as:

$$\max_{Q_i} (1 - \theta)(M_i - Q_i + Q_i d + w_i L_i) + \theta(M_i - Q_i) - v_i(L_i) \quad (1.120)$$

s.t.

$$M_i - Q_i \geq c_{0,i} \quad (1.121)$$

The derivative of the above with respect to  $Q_i$  is:

$$(1 - \theta)d - 1 \quad (1.122)$$

Note that, in equilibrium, the dividends per share are given by:

$$d = \frac{M_i}{Q_i^0} \quad (1.123)$$

**Assumption 3** *The value of  $\theta$  is sufficiently small so that, for any realization of  $M_i$  and  $c_{0,i}$ , the derivative of equation 1.120 with respect to  $Q_i$  is positive at  $Q_i^0 = M_i - c_{0,i}$ :*

$$(1 - \theta) \frac{M_i}{M_i - c_{0,i}} - 1 > 0 \quad (1.124)$$

**Result 1** *Under Assumption 3, the optimal portfolio decision is  $M_{h,i} = c_{0,i}$ ,  $Q_i = M_i - c_{0,i}$ .*

**Proof:** Under Assumption 3, the constraint in equation 1.121 is binding. Thus,

$$M_i - Q_i = c_{0,i}.$$

## 1.10.2 Liquidity Supply in the Integrated Economy

I assume that there is a single currency, so both labor and consumption can be paid for in either domestic or foreign money. For simplicity, I assume that agents consume foreign and domestic goods proportional to their shares in output (so that, regardless of the realizations of money supplies, a unit of domestic output has the same probability of being consumed as a unit of foreign output).

As in the closed economy, agents choose between keeping money in the safe ( $M_{h,i}$ ) and buying stocks ( $Q_i$ ). In the integrated economy, a stock is a claim on the sales revenues of the global economy. After portfolio decisions are made, the global financial system distributes liquidity between foreign and domestic projects in an output maximizing way<sup>20</sup>.

**Assumption 4** *Shocks to the ability of the productive sector to make monetary transfers are i.i.d. across regions.*

### Equilibrium

**The portfolio decision.** Because there is a positive probability that both economies suffer a simultaneous collapse (an event that happens with probability  $\theta^2$ ), similarly to the closed economy case the agent will reserve at least enough money to finance  $c_{0,i}$  units of consumption:

$$M_{h,i} \geq c_{0,i} \tag{1.125}$$

Denote by  $d$  the realized dividend per share of the global productive sector, and let  $Q_i^0$  be the equilibrium level of liquidity supplied by country  $i$  which the individual

---

<sup>20</sup>In this model, it is implicitly assumed that consumption is always less than output, and that there are some units of output which are ex-post “wasted”. A natural question is therefore why the global financial system allocates liquidity between foreign and domestic projects in an output maximizing way. If we think of the global financial system as a monopoly, this indeed need not be the case; however, a more competitive structure (for example, one in which the financial system is composed of many small banks competing for liquidity) would deliver this result, as the expected real value of a unit of produced output is positive.

agent takes as given. Similarly to the closed economy, agents solve:

$$\max_{Q_i, M_{h,i}} c_i - v_i(L_i) \quad (1.126)$$

s.t.:

$$M_i = Q_i + M_{h,i} \quad (1.127)$$

$$c_i = \begin{cases} M_{h,i} + Q_i d + w_i L_i & \text{with prob. } (1 - \theta); \\ M_{h,i} + Q_i d & \text{with prob. } (1 - \theta)\theta; \\ M_{h,i} & \text{with prob. } \theta^2. \end{cases} \quad (1.128)$$

$$M_{h,i} \geq c_{0,i} \quad (1.129)$$

Substituting in the first constraint, this problem can be rewritten as:

$$\max_{Q_i} E(Q_i d) + (1 - \theta)w_i L_i + (M_i - Q_i) - v_i(L_i) \quad (1.130)$$

s.t.

$$M_i - Q_i \geq c_{0,i} \quad (1.131)$$

The derivative of the above with respect to  $Q_i$  is:

$$E(d) - 1 \quad (1.132)$$

To calculate  $E(d)$ , note that positive dividends are paid unless both economies suffer a collapse. If neither suffers a collapse (an event which occurs with probability  $(1 - \theta)^2$ ), dividends per share are  $d = \frac{M_d + M_{em}}{Q_d^0 + Q_{em}^0}$ . If economy  $i$  suffers a collapse but economy  $j$  doesn't suffer a collapse (events which occur with probability  $(1 - \theta)\theta$  each), the dividend per share is positive, as consumers in region  $i$  receive some dividends from their stock holdings in region  $j$ . Expected dividends per share are therefore bounded from below by:

$$E(d) > (1 - \theta)^2 \frac{M_d + M_{em}}{Q_d^0 + Q_{em}^0} \quad (1.133)$$

**Assumption 5** *The value of  $\theta$  is sufficiently small so that for any realization of  $M_i$*



and  $c_{0,i}$ , the following inequality holds:

$$(1 - \theta)^2 \frac{M_d + M_{em}}{(M_d - c_{0,d}) + (M_{em} - c_{0,em})} - 1 > 0 \quad (1.134)$$

**Result 2** Under Assumption 5, the optimal portfolio decision is  $M_{h,i} = c_{0,i}$ ,  $Q_i = M_i - c_{0,i}$ .

**Proof:** Under Assumption 5, the derivative of equation 1.130 with respect to  $Q_i$  is positive at  $Q_i^0 = M_i - c_{0,i}$ :

$$E(d) - 1 > (1 - \theta)^2 \frac{M_d + M_{em}}{Q_d^0 + Q_{em}^0} - 1 = (1 - \theta)^2 \frac{M_d + M_{em}}{(M_d - c_{0,d}) + (M_{em} - c_{0,em})} - 1 > 0 \quad (1.135)$$

Thus, the constraint in equation 1.131 is binding, so  $M_i - Q_i = c_{0,i}$ .

**Corollary 3** Both under autarky and under financial integration, the liquidity supply of country  $i$  is given by  $Q_i = M_i - c_{0,i}$ .

**Assumption 6** The money supply,  $M_i$ , and the risk aversion parameter,  $c_{0,i}$ , are perfectly correlated across regions and follow time invariant distributions.

Since equilibrium liquidity supply is  $Q_i = M_i - c_{0,i}$  both in the integrated economy and under autarky, it follows trivially that the assumptions I make on the distribution of liquidity supply are satisfied.



## Chapter 2

# The Inefficiency of Financial Intermediation in General Equilibrium

### Abstract

In the presence of liquidity constraints, there are rents from supplying liquidity to constrained entrepreneurs. In partial equilibrium, when the price of inputs is fixed in terms of liquidity, a financial system facilitates the efficient allocation of resources by relaxing liquidity constraints. However, in general equilibrium, the presence of a financial sector has two adverse implications: first, intermediation activities absorb productive resources, reducing the amount of inputs employed by the productive sector. Second, financial intermediation bids up the price of inputs in terms of liquidity, increasing the economy's dependence on the financial sector, which is a source of crises. Consequently, the presence of a financial sector may reduce equilibrium welfare. I show that an optimal policy is to tilt the tradeoff between production and liquidity hoarding in favor of liquidity hoarding. The optimal policy serves both to relax liquidity constraints and to crowd out the financial sector.

**JEL Classification:** E44, E6, G01, G28, H21

**Keywords:** Costs of the financial sector, financial intermediation, output tax, Friedman rule, financial crises

## 2.1 Introduction

The recent financial crisis resurfaced the concern that too many productive resources are being absorbed by the financial sector, and that the vulnerability of the real sector to “mistakes” by the financial sector is too large<sup>1</sup>. However, intermediation theory suggests that financial intermediation improves efficiency by improving the allocation of productive resources and reducing liquidity constraints<sup>2</sup>. In this paper I reconcile these views by demonstrating that while financial intermediation may be vital in partial equilibrium, in general equilibrium, when the price of inputs is adjustable, financial intermediation is potentially a wasteful use of productive resources.

I propose that the government can reduce the need for intermediation by tilting the tradeoff between productive and unproductive saving in favor of unproductive saving. This can be done either by taxing production or by subsidizing liquidity hoarding. Increasing the benefits of liquidity hoarding relative to productive investment will achieve two things: first, it will lower the demand for capital and hence its equilibrium price, which will enable constrained firms to produce more with the limited funds available to them. Second, it will lower the returns to financial intermediation; the financial sector will shrink, freeing up resources for the productive sector.

The proposed policy favors government intervention over private financial intermediation. Despite this, I argue that it is closer in spirit to the free market paradigm. The system of financial intermediation resembles a privately-run central planning system, in which the allocation of the means of production is determined by a group of “experts” rather than by market forces. The “invisible hand” is replaced by bankers trying to evaluate market conditions to determine where resources are best allocated. Of course, unlike central planning systems, the opportunities for bankers to make large private profits help insure that their decisions are not tainted by political or other personal considerations; in this sense, financial intermediation is clearly preferable to central planning. However, experience suggests that even the most well-motivated central planning system cannot mimic the outcome of perfectly competitive markets,

---

<sup>1</sup>See op-eds by Friedman [2009] and Volcker [2010].

<sup>2</sup>See Gorton and Winton [2003] for a survey of the literature on financial intermediation.

in which prices and production decisions reflect an aggregation of beliefs about supply and demand, including an entrepreneur's personal knowledge of his own skills and the demand for his product. In order to achieve a profit-maximizing resource allocation, financial intermediaries must engage in costly information acquisition, demeaning monitoring, and exercise harsh repercussions in case of default. By instituting a simple tax on output that changes the returns to production, liquidity constraints are relaxed, and production decisions may be taken freely and independently by entrepreneurs, without relying on intermediation. The "invisible hand" comes back into play and allows the allocation of resources to be determined according to traditional supply and demand forces.

The mechanism in this paper is conceptually related to the Friedman Rule (Friedman [1969]). In a Friedman-rule monetary economy, cash-in-advance constraints may be binding in equilibrium, while in the socially efficient allocation they never bind. In a monetary framework, the inefficiency of equilibrium stems both from missed consumption opportunities and from a waste of resources spent on "trips to the bank". The government can help alleviate these inefficiencies by subsidizing money holding. The model in this paper is somewhat analogous: in equilibrium, entrepreneurs' liquidity constraints are binding, while in the socially efficient allocation they are not. Similar to the monetary framework, there are two sources of inefficiency in equilibrium: first, the allocation of production inputs is inefficient. This is somewhat analogous to the missed consumption opportunities in Friedman's economy: productive entrepreneurs miss out on high-return production opportunities because they do not have enough liquidity. The second source of inefficiency is the wasteful resources spent on financial intermediation. This is analogous to the wasteful resources spent on trips to the bank in the monetary framework. The proposed subsidy on liquidity hoarding is similar in spirit to the proposition to subsidize money holding in a Friedman rule economy.

This paper is related to the literature on the government's role in alleviating liquidity constraints, such as Aiyagari and McGrattan [1998], Holmstrm and Tirole [1998], Gorton and Huang [2004], and Eden [2010]. In contrast to this paper, the role

of the government in the papers mentioned above is to directly provide loans to alleviate liquidity constraints, rather than manipulate prices such that the equilibrium allocation of capital is efficient. Most closely related to this paper is Eden [2010]. B. Eden argues that the government has a technological advantage over the financial sector in providing loans. The technological advantage results from the government's superior ability to verify income, as well as from its preference towards redistribution which allows it to view default as a transfer. The argument for an output tax presented here shares the view that the government should play a more prominent role in alleviating liquidity constraints and crowd out some of the activity of the financial sector. However, unlike Eden [2010], the focus here is on the productive sector, and the government's advantage stems from its ability to change the relative price of investment in a way which allows traditional equilibrium forces to operate. Unlike Eden [2010], the mechanism through which the government alleviates liquidity constraints is different in nature from the mechanism employed by the financial sector; it does not provide the same service more cheaply but rather manipulates equilibrium prices so that financial services are no longer needed.

This paper also relates to the literature on optimal capital taxation in an environment with borrowing constraints. Aiyagari [1995] demonstrates that in an environment with borrowing constraints, households have an incentive to over-invest in capital because of a precautionary saving motive. He concludes that a positive tax on capital is welfare-improving. The important difference between Aiyagari [1995] and the setup in this paper is that Aiyagari assumes that firms are never credit constrained, so that the rental rate of capital is equated with its marginal product. Here, I assume that firms are credit constrained; as a consequence, the rental rate of capital is depressed (even in the absence of any policy intervention, but more so under the optimal policy). There is therefore an under-incentive to accumulate capital, and the optimal policy is to subsidize capital supply.

A positive gap between the marginal product of capital and its market price is possible in this model because entrepreneurs cannot accumulate capital for their own production needs and must purchase it through a market. Conceptually, in a

growing economy in which future output is not pledgable, this friction would also prevent entrepreneurs from accumulating sufficient funds to overcome their liquidity constraints (as in Aiyagari [1994]). This is because the aggregate payment to capital is always bounded by current output, which may not suffice to account for future returns to capital. The friction preventing entrepreneurs from accumulating capital for self-production represents a realistic type of “monetary” friction: in reality, production requires a variety of capital inputs, where there are some increasing returns in the production of each input. The investment in capital required for an entrepreneur to be self-sufficient is far too expensive. The entrepreneur therefore prefers to purchase inputs through a market.

This paper is also related to the literature on the optimal transaction tax and the regulation of the financial sector, such as Jacklin [1987] or, more recently, Scheuer [2009]. The role of the financial sector emphasized in this paper is different from the one emphasized in the papers listed above, as the focus is on financial intermediation in a world with no uncertainty rather than on trade in state-contingent claims. Farhi et al. [2009] focus similarly on the role of the financial sector in providing liquidity, and argue that the higher return to long run projects (as in Diamond and Dybvig [1983]) may cause aggregate liquidity supply to be too low. This provides further foundation for the view that liquidity constraints are in some sense inevitable even in the long-run. This also raises an additional potential benefit of the policy proposed here: with a sufficiently high output tax, the first best level of output can be achieved with *any* level of aggregate liquidity, as the price of capital inputs adjusts appropriately. This allows the economy to invest in more high yield long term projects, as there is no need for costly supply of liquidity.

The rest of this paper is organized as follows. In section 2.2, I lay out a stylized model of liquidity constraints in an economy with no financial intermediation. I show that equilibrium welfare can be improved by instituting a tax on production or, equivalently, a subsidy on liquidity hoarding. In section 2.3, I enrich the model by introducing a financial intermediation technology, and show that the financial sector is eliminated under the optimal policy. In section 2.4 I show how these results

generalize to richer environments. In section 2.5, I discuss realistic concerns regarding the proposed policy. In section 3.8 I conclude.

## 2.2 Benchmark: an Economy with no Financial Intermediation

Prior to introducing financial intermediation, I demonstrate that a tax on output can improve resource allocation by discouraging inefficient entrepreneurs from self-financing. In section 2.3 I enrich the model by allowing for a costly intermediation technology.

There is a measure 1 of *entrepreneurs* indexed  $x \in [0, 1]$ , and a measure 1 of *capital owners*. Capital is the single productive input which is at a fixed supply  $\bar{K}$ . Capital is owned initially by capital owners, who may sell their capital to entrepreneurs (the scrap value of capital is assumed to be 0, so the only way that capital owners can consume is by selling their capital to entrepreneurs).

To abstract from distributional concerns, I assume that agents' consumption utility is linear and that ex-ante agents do not know their identity, including whether they are capital owners or entrepreneurs. Given this assumption, welfare is given simply by the level of output.

Entrepreneur  $x$  is endowed with the following  $AK$  production technology:

$$Y(x) = A(x)K(x) \tag{2.1}$$

Where  $K(x)$  is the capital employed by entrepreneur  $x$ , and the productivity parameter,  $A(x)$ , is decreasing in  $x$ . For simplicity, I assume that there are only two types of entrepreneurs, so that  $A(x)$  takes the following form:

$$A(x) = \begin{cases} \bar{A} & \text{if } x \in [0, \frac{1}{2}]; \\ \underline{A} & \text{otherwise.} \end{cases} \tag{2.2}$$



Where:

$$\bar{A} > \underline{A} \geq 0 \tag{2.3}$$

Entrepreneurs  $[0, \frac{1}{2}]$  will be referred to as *productive entrepreneurs*, and entrepreneurs  $(\frac{1}{2}, 1]$  will be referred to as *unproductive entrepreneurs*.

To summarize, the timeline of the model is as follows:

- $t = 1$ : Entrepreneurs buy capital from capital owners.
- $t = 2$ : Entrepreneurs produce and consume.

**The efficient economy.** The first best allocation of in this economy is characterized by the following lemma:

**Lemma 8** *The welfare maximizing allocation of capital is having the productive entrepreneurs employ the entire capital stock, that is:  $K(x) = 0$  for  $x \in (\frac{1}{2}, 1]$ , and  $\int_0^{\frac{1}{2}} K(x)dx = \bar{K}$ .*

The proof of the lemma stems from the fact that production is most efficient when carried out by productive entrepreneurs, and there are no decreasing returns to capital.

**Lemma 9** *In the absence of any further frictions, the welfare maximizing allocation is achieved in any equilibrium.*

**Proof:** Consider the equilibrium determination of the rental rate of capital,  $R$ . In equilibrium,  $R = \bar{A}$ : if  $R < \bar{A}$ , entrepreneur 0 has strictly positive demand for capital. If  $R > \bar{A}$ , aggregate demand for capital is 0. Capital market clearing therefore implies  $R = \bar{A}$ . At  $R = \bar{A}$ , productive entrepreneurs are indifferent whether or not to buy capital, whereas unproductive entrepreneurs strictly do not want to buy capital. Thus, it is concluded that in equilibrium the entire capital stock is employed by productive entrepreneurs.

Note that this equilibrium can be implemented under various assumptions regarding the resources available to entrepreneurs before production takes place. For example, if output is pledgable, entrepreneurs can buy capital with claims on future output. Alternatively, this equilibrium can be implemented if it is assumed that entrepreneurs have a large stock of initial funds which can be used to finance the purchase of inputs (the exact condition is that productive entrepreneurs have at least  $\bar{A}\bar{K}$  initial funds, which is the payment to capital in the competitive equilibrium).

**Liquidity constraints.** I modify the model to allow for liquidity constraints. I assume that output is not pledgable; capital must be bought with current consumption goods<sup>3</sup>. Initial endowments of consumption goods are identical across entrepreneurs (and 0 for capital owners): each entrepreneur is born with  $Q$  units of consumption goods, which will be referred to as the entrepreneur's *liquidity*. Liquidity can be stored across periods, so that entrepreneurs can choose to store their liquidity and consume it later.

I assume that the amount of liquidity in the hand of productive entrepreneurs is not enough to finance the payment to capital in the unconstrained equilibrium:

$$\frac{1}{2}Q < \bar{A}\bar{K} \tag{2.4}$$

Entrepreneurs can choose between consuming their endowment and using it to buy capital. The act of storing the liquidity endowment (and consuming it later) will be referred to as *liquidity hoarding*. This should be thought of as using liquidity for other purposes outside of productive investment. For example, this can include liquidity services for consumption, or saving in non-productive assets such as treasuries or consumption loans. The act of using the liquidity endowment to purchase capital will be referred to as *self-financing*.

In equilibrium, entrepreneurs with sufficiently high productivity will self-finance, as they always receive rents from production.

---

<sup>3</sup>The assumption that future output is not pledgable is in the spirit of Kehoe and Levine [1993], Holmstrom and Tirole [1997], Kiyotaki and Moore [1997] or Caballero and Krishnamurthy [2001].

**Lemma 10** *There are some unproductive entrepreneurs who choose to self-finance in equilibrium if and only if the following condition holds:*

$$\underline{A} > \frac{\frac{1}{2}Q}{\bar{K}} \quad (2.5)$$

The proof is in the appendix.

**Corollary 4** *If the condition in equation 2.5 holds, then the equilibrium allocation of capital is inferior to the first best. Otherwise, output is at its first best level even though the liquidity constraint is binding for productive entrepreneurs.*

**Proof:** By Lemma 10, if the condition in equation 2.5 holds, some of the capital stock is employed by unproductive entrepreneurs. This is suboptimal because the optimal allocation of capital is to have the entire capital stock employed by productive entrepreneurs. Similarly, if the condition in equation 2.5 is violated, unproductive entrepreneurs choose not to self finance. The entire capital stock is therefore employed by productive entrepreneurs, and the first best level of output is achieved.

The case in which unproductive entrepreneurs employ capital resembles the steady state outcome in Kiyotaki and Moore [1997]. In their model, the return to capital for “farmers” (productive entrepreneurs) is strictly higher than the price of employing capital. This gap cannot be bridged because farmers cannot borrow against future returns. As a result, some of the capital is allocated to “gatherers”, who are less efficient in production but are able to pledge their entire output.

### 2.2.1 Optimal Policy

An important implication of the model in this paper is that taxing production and subsidizing liquidity hoarding can improve the allocation of capital, thus carrying out the role typically designated for financial intermediation. In this section, I demonstrate that the government can improve the allocation of resources without using any form of financial intermediation. The efficient allocation of resources is done without

any transfers of liquidity from one entrepreneur to another. In other words, with liquidity constraints, the equilibrium is constrained inefficient as in Davila et al. [2007]: the government could theoretically improve allocations simply by prescribing different saving behavior. Under an optimal policy, taxes may be redistributed lump sum to entrepreneurs, but this is done only for the purpose of increasing consumption. The efficient allocation of resources does not rely on the “intermediation” of liquidity through the government.<sup>4</sup>

I assume that the government can set a tax  $\tau$  on output: an entrepreneur must pay the government a fraction  $\tau$  of any amount of output he produces. Additionally, the government can set a subsidy  $\epsilon$  on liquidity hoarding: a unit of stored liquidity earns interest at a rate of  $\epsilon$ . It will be shown that both tax instruments are equivalent. Importantly, both the output tax and the hoarding subsidy lower the relative return to production for all agents. In particular, agents at the margin who are roughly indifferent between self financing and liquidity hoarding will switch over to strictly preferring liquidity hoarding. This frees up productive resources which are then employed by more productive entrepreneurs, who still find it optimal to produce despite the relatively higher returns to liquidity hoarding.

Denote by  $Y(\tau, \epsilon)$  the output produced in equilibrium when the output tax is  $\tau$  and the hoarding subsidy is  $\epsilon$ , and by  $Y^{FB}$  the first best level of output (which is achieved when all productive inputs are employed by productive entrepreneurs).

Tables 2.1, 2.2 and 2.3 summarize the notation used in this paper (the variables  $h$  and  $\mu$  will be introduced later in the text).

Table 2.1: Policy instruments

Variable	Notation
Output tax	$\tau$
Hoarding subsidy	$\epsilon$

I assume that all taxes are collected (or distributed) **only after production takes place**. This is important because otherwise the government can bypass the

---

<sup>4</sup>This is an important difference with Eden [2010], in which the government is able to provide the same service as the financial sector only more cheaply.

Table 2.2: Primitives

Variable	Notation
Liquidity endowment	$Q$
Fraction of capital required for intermediation	$\mu$
Capital stock	$\bar{K}$
Productivity of productive ent.	$\bar{A}$
Productivity of unproductive ent.	$\underline{A}$

Table 2.3: Equilibrium objects

Variable	Notation
Price of capital	$R$
Fraction of aggregate liquidity hoarded	$h$
Output	$Y$

liquidity constraints directly by subsidizing liquidity and effectively lending to constrained entrepreneurs.

In addition to the tax instruments listed above, it is possible to allow for a lump-sum transfer (post production), through which the government can distribute tax revenues among consumers. However, in the formulation above it is not necessary as the government can redistribute tax revenues as hoarding subsidies ( $\epsilon$ ). It turns out that in this framework, for any given tax, this form of redistribution is output-maximizing as it further tilts the tradeoff between production and liquidity hoarding in favor of liquidity hoarding. In section 2.5.1 I show that given an elastic supply of capital, the output-maximizing form of redistribution may also include a subsidy on capital supply.

Let  $\tau^* = \tau^*(\epsilon)$  and  $\epsilon^* = \epsilon^*(\tau)$  be given by:

$$(1 - \tau^*)\underline{A}\frac{1}{R^*} = 1 + \epsilon \quad (2.6)$$

$$(1 - \tau)\underline{A}\frac{1}{R^*} = 1 + \epsilon^* \quad (2.7)$$

Where  $R^*$  is the price of liquidity under which the productive entrepreneurs have enough liquidity to employ the entire capital stock:

$$\frac{1}{2}Q = R^* \bar{K} \Rightarrow R^* = \frac{Q}{2\bar{K}} \quad (2.8)$$

The left hand side of equations 2.6 and 2.7 represent the returns to self-financing by unproductive entrepreneurs given the price  $R^*$ . The right hand side represents the return to liquidity hoarding. It follows that if  $\tau = \tau^*$ , or  $\epsilon = \epsilon^*$ , unproductive entrepreneurs are indifferent between self-financing and liquidity hoarding; in particular, they are willing to hoard liquidity and refrain from production.

**Proposition 7** *In an environment without financial intermediation:*

1. *Output monotonically approaches first best as  $\tau$  approaches  $\tau^*$ :*

- (a)  $\frac{\partial Y(\tau, \epsilon)}{\partial \tau} \geq 0$ , and
- (b)  $Y(\tau^*(\epsilon), \epsilon) = Y^{FB}$

2. *Output monotonically approaches first best as  $\epsilon$  approaches  $\epsilon^*$ :*

- (a)  $\frac{\partial Y(\tau, \epsilon)}{\partial \epsilon} \geq 0$ , and
- (b)  $Y(\tau, \epsilon^*(\tau)) = Y^{FB}$

The intuition behind Proposition 7 is as follows. The misallocation of capital results from the fact that productive entrepreneurs do not have sufficient funds to employ the entire capital stock. Generally speaking, financial intermediation alleviates this problem by increasing the funds in the hands of productive entrepreneurs. The production tax takes the dual approach to intermediation: rather than increasing the funds in the hands of entrepreneurs, it lowers the equilibrium price of capital, enabling constrained entrepreneurs to purchase more capital with the funds available to them. The equilibrium feedback of the production tax is that less efficient users of capital find it less desirable to produce, thus decreasing the demand for capital. As a result, the price of capital declines, and productive entrepreneurs who still find it optimal to produce can afford to purchase more capital.

The assumption that the capital stock is fixed is crucial for this result. The standard argument states that a tax on output lowers output, because it decreases

the incentives to supply inputs. This channel is shut down here by assuming that the capital stock is fixed. In section 2.5.1 I show that this concern can be addressed by issuing a subsidy on capital supply.

To prove the proposition formally, denote by  $h$  the fraction of aggregate liquidity being hoarded. The price of capital is given by the following market clearing condition, that equates the supply of capital with the demand for capital:

$$R\bar{K} = Q(1 - h) \Rightarrow R = \frac{Q(1 - h)}{\bar{K}} \quad (2.9)$$

Entrepreneur  $x$  chooses to employ capital if and only if:

$$(1 - \tau)A(x)\frac{1}{R} > 1 + \epsilon \Leftrightarrow (1 - \tau)A(x)\frac{\bar{K}}{Q(1 - h)} > 1 + \epsilon \quad (2.10)$$

Since equilibrium requires capital market clearing, it must be the case that prices are such that productive entrepreneurs, for whom the return to capital is the highest, are always willing to employ capital. The condition in equation 2.10 must therefore hold for  $A(x) = \bar{A}$ .

If  $h = 0$ , all entrepreneurs strictly prefer to refrain from liquidity hoarding; the equilibrium will not change with a small increase in  $\tau$  or in  $\epsilon$ . If  $\frac{1}{2} > h > 0$ , unproductive entrepreneurs must be indifferent between liquidity hoarding and self-financing. The level of liquidity hoarding is determined by the following equilibrium condition:

$$(1 - \tau)\bar{A}\frac{\bar{K}}{Q(1 - h)} = 1 + \epsilon \Rightarrow (1 - h) = (1 - \tau)\bar{A}\frac{\bar{K}}{Q(1 + \epsilon)} \quad (2.11)$$

From the above formulation, it is evident that  $h$  increases with  $\tau$  and with  $\epsilon$ .

Output is given by:

$$Y(\tau, \epsilon) = \frac{1}{2}\bar{A}\frac{\frac{1}{2}}{1 - h}\bar{K} + \frac{1}{2}\bar{A}\frac{1 - h - \frac{1}{2}}{1 - h}\bar{K} \quad (2.12)$$

The first term is the output produced by productive entrepreneurs. The share of capital that they employ is equal to the share of their liquidity endowments in total liquidity used for productive purposes:  $\frac{\frac{1}{2}}{1 - h}$ . The second term is the output produced

by unproductive entrepreneurs who choose not to hoard liquidity. The share of the capital that they employ is similarly given by their share of liquidity relative to total liquidity used for productive purposes,  $\frac{1-h-\frac{1}{2}}{1-h}$ .

From equation 2.12, it is easy to see that output is increasing in  $\tau$  and  $\epsilon$ . Moreover,

$$\lim_{\tau \rightarrow \tau^*} h = \frac{1}{2} \quad (2.13)$$

$$\lim_{\epsilon \rightarrow \epsilon^*} h = \frac{1}{2} \quad (2.14)$$

At  $h = \frac{1}{2}$ , the allocation of capital is efficient, as productive entrepreneurs employ the entire capital stock.

Note that the production tax improves efficiency by increasing the set of entrepreneurs who find it optimal not to produce but to hoard liquidity. At first glance, this may seem paradoxical: compared to the efficient economy, the problem in the liquidity-constrained economy is that there is not enough liquidity in the hands of the productive sector. Yet, to improve efficiency, we are asking agents to refrain from using their liquidity for production, lowering the amount of liquidity in the system even further. The key to understanding why this works is to realize the sense in which there is *not enough* liquidity in the liquidity-constrained economy: there is not enough liquidity to allow for input prices to account for their returns. As demonstrated by the efficient economy case, if liquidity was sufficiently abundant, the price of capital would be bid up to the marginal product of the most efficient entrepreneur. Thus, only the productive entrepreneurs would find it optimal to produce, and resources would be allocated efficiently. The problem with not having *enough* liquidity is that input prices are depressed, so unproductive entrepreneurs find it optimal to produce. The direct way to improve welfare is to discourage unproductive entrepreneurs from producing.

Note that the reason that liquidity hoarding should be encouraged is that in *general* equilibrium, the price of inputs adjusts so that productive entrepreneurs can hire more inputs with the limited funds available to them. The negative association of liquidity hoarding with crisis amplification is consistent in partial equilibrium, when



the prices of inputs are fixed<sup>5</sup>. I address this issue in section 2.5.2, in which I consider an environment with sticky prices.

## 2.3 An Economy with Financial Intermediation

The gap between the market price of capital and the marginal return to capital is fertile grounds for the emergence of a financial intermediation system. Financial intermediation transmits liquidity from unproductive entrepreneurs to productive entrepreneurs at some cost. This could be a cost of monitoring (as in Townsend [1979], Diamond [1984], and Williamson [1986]), of acquiring information (as in Leland and Pyle [1977], Campbell and Kracaw [1980], and Boyd and Prescott [1986]), or of transacting (as in Benston and Smith [1976]). Typically, in models of financial intermediation, banks reduce the cost associated with reallocating liquidity from unproductive to productive entrepreneurs; however, there are typically still resources which need to be spent in order to overcome the initial friction preventing entrepreneurs from borrowing from lenders directly. Realistically, the large amount of productive resources spent on financial intermediation is a source of concern<sup>6</sup>.

I proceed in modifying the framework to allow for financial intermediation. I assume that agents have an option to activate an intermediation technology. If agent  $y$  wants to make use of the technology available to entrepreneur  $x$ , he may do so indirectly by buying capital and allowing entrepreneur  $x$  to use it. The agents agree on a repayment rate,  $r$  per unit of capital, to be paid after production takes place. For simplicity, I assume that in order to activate this intermediation technology, agent  $y$  must employ  $\frac{1-\mu}{\mu}$  units of capital per unit of capital deposited with agent  $x$ . This can be thought of as the cost of monitoring, contracting, etc<sup>7</sup>. If agent  $x$  uses capital

---

<sup>5</sup>For examples of the negative implications of liquidity hoarding in crises situations, see Caballero and Krishnamurthy [2008] or Caballero and Simsek [2010].

<sup>6</sup>The profits of the financial sector in the US range between 4-8% of GDP. This suggests that the resources spent on intermediation are non-negligible.

<sup>7</sup>In many models of financial intermediation, there is an element of increasing returns: the more capital intermediated the less the cost of intermediation per unit of capital. I abstract from this and assume that a constant fraction of capital is absorbed by the financial sector. It is easy to show that the results in this section generalize as long as the total amount of resources spent on intermediation increases with the amount of intermediation.

owned by agent  $y$ , I say that agent  $x$  uses *intermediated capital*.

To summarize, the timeline is modified as follows:

- $t = 1$ : Entrepreneurs buy capital from capital owners. Entrepreneurs choose among the following options:
  - **Self financing**: entrepreneurs buy capital and use it to activate their own production technology.
  - **Liquidity hoarding**: entrepreneurs do not buy capital but store their liquidity endowment for later consumption.
  - **Intermediation**: entrepreneurs buy capital and deposit it with the entrepreneur of their choice. A fraction  $\mu$  of the capital they purchase is used for activating the intermediation technology.
- $t = 2$ : Production and taxation. Entrepreneurs consume the net-of-tax output that they produced and the liquidity that they have stored. Capital owners consume the sales revenues from capital.

**Lemma 11** *In equilibrium, only the productive entrepreneurs,  $x \in [0, \frac{1}{2}]$ , use intermediated capital. The repayment rate is the entire output produced by the intermediated capital ( $r = \bar{A}$ ).*

The proof is in the appendix.

The assumption that using financial intermediation is costly while taxation is free requires some elaboration. In this model, the superiority of an output tax stems from the assumption that tax collection is cheaper than financial intermediation. I dedicate section 2.5.3 to elaborate on this assumption.

### 2.3.1 Is Financial Intermediation Welfare-Improving?

Prior to considering optimal policy in an environment with financial intermediation, I address the question of whether or not the existence of an intermediation technology is welfare-improving in an environment with no government policy. It will be shown

later that the optimal policy is such that there is no financial intermediation in equilibrium.

It turns out that whether the presence of an intermediation technology improves equilibrium welfare depends on the primary alternative for financial intermediation. If financial intermediation comes mainly as a substitute for liquidity hoarding or non-productive savings, then it reduces welfare in equilibrium. It may improve equilibrium welfare if it comes mainly as a substitute for self-financing. I demonstrate this principle with two examples:

**Example 1:** Assume that the productivity distribution is given by  $\bar{A} > 0$  and  $\underline{A} = 0$ .

As long as the price of inputs is positive, unproductive entrepreneurs will never choose to produce. Since the productivity of all unproductive entrepreneurs is high, in equilibrium they will all choose to produce.

Consider first the equilibrium with no financial intermediation. In this case, unproductive entrepreneurs choose to hoard liquidity. The entire capital stock is employed by productive entrepreneurs, so the total output in the economy with no intermediation ( $ni$ ) is equal to the first best level of output:

$$Y^{ni} = \bar{A}\bar{K} = Y^{FB} \quad (2.15)$$

Consider now the equilibrium with financial intermediation. The high profits generated from capital in the hands of productive entrepreneurs provide an incentive for unproductive entrepreneurs to employ the intermediation technology. Thus, all of the liquidity in the economy is used to hire capital (either for intermediation or for production). The equilibrium is inferior to the no intermediation case, because some of the capital stock is used for intermediation rather than for production. Specifically, only a fraction  $\frac{1}{2} + (1 - \mu)\frac{1}{2}$  of the capital stock is employed by the productive sector, while a fraction  $\frac{1}{2}\mu$  is employed in intermediation. Equilibrium output with financial intermediation ( $wi$ ) is lower

than equilibrium output without financial intermediation:

$$Y^{wi} = \bar{A}\left(\frac{1}{2} + (1 - \mu)\frac{1}{2}\right)\bar{K} < A\bar{K} = Y^{FB} \quad (2.16)$$

By absorbing productive resources, the existence of a financial intermediation technology reduces equilibrium welfare. The key is that financial intermediation does not improve the allocation of capital, but only transfers some of the production revenues from productive to unproductive entrepreneurs. The activation of the intermediation technology is costly from a social perspective, but emerges in equilibrium because it enables a transfer of wealth to those who activate it.

**Example 2:** Let  $A \gg 0$ , and assume that the productivity distribution is given by  $\bar{A} = A$  and  $\underline{A} = \lambda A$  (where  $\lambda < 1$ ). In the absence of financial intermediation, if the liquidity constraint is sufficiently tight ( $Q$  is sufficiently small), the equilibrium price of capital is sufficiently depressed that profits from self financing are positive even for relatively unproductive entrepreneurs. Thus, equilibrium output in the no intermediation case ( $ni$ ) is given by:

$$Y^{ni} = \frac{1}{2}A\bar{K} + \frac{1}{2}\lambda A\bar{K} \quad (2.17)$$

If the capital requirement for the financial intermediation technology is not too large, financial intermediation can improve equilibrium welfare. Importantly, in this case the decision to use the intermediation technology is efficient: financial intermediation will be employed in equilibrium only if it improves welfare. To see this, note that unproductive entrepreneurs will choose to employ the intermediation technology only if  $(1 - \mu) > \lambda$ . This is also the condition under which intermediation improves welfare. If the intermediation technology is employed, output is given by:

$$Y^{wi} = \frac{1}{2}A\bar{K} + \frac{1}{2}(1 - \mu)A\bar{K} > Y^{ni} \quad (2.18)$$

Where the inequality stems from the fact that the intermediation technology is employed in equilibrium only if  $(1 - \mu) > \lambda$ .

The examples above demonstrate that the existence of an intermediation technology has a mixed effect on equilibrium welfare. The intermediation technology can help improve the allocation of capital by providing a more productive alternative for unproductive entrepreneurs who choose to self-finance. However, the employment of the intermediation technology is unjustified from a welfare perspective when the alternative to intermediation is unproductive savings. The employment of the intermediation technology is wasteful as it absorbs productive resources and does not improve the equilibrium allocation of capital.

There have been recent proposals to institute a “transaction tax”, which are aimed at discouraging the usage of financial intermediation. This model suggest that the welfare implications of such policies depend crucially on the most attractive alternative use of funds: an intermediation tax may be beneficial if it encourages liquidity hoarding, but may be harmful if it encourages self-financing. However, there is a strong argument for eliminating subsidies for intermediation. It can be argued that the inability of the government to commit not to bail out the financial sector during a crisis prevents agents from realizing the entire social cost of intermediation; hence, financial intermediation is effectively subsidized. In section 2.5.2, I demonstrate that during a financial crisis, it is indeed optimal for the government to bail out the financial sector. This effective subsidy may therefore be corrected by a “normal time” tax on financial intermediation. However, it will be shown that under the optimal policy there is no financial intermediation.

An alternative specification of the cost of intermediation is to allow for the possibility of a financial crisis: after prices are set, there is some probability  $\theta$  that the financial sector collapses and intermediation is no longer possible. Consider the specification of example 1, in which  $\bar{A} > 0$  and  $\underline{A} = 0$ . In the absence of financial intermediation, output is at its first best level. With financial intermediation,

$$Y^{wi} = \begin{cases} \bar{A}(\frac{1}{2} + (1 - \mu)\frac{1}{2})\bar{K} & \text{with prob. } 1 - \theta; \\ \bar{A}(\frac{1}{2} + 0)\bar{K} & \text{with prob. } \theta. \end{cases} \quad (2.19)$$

Note that with the possibility of a financial crisis, the presence of a financial sector reduces equilibrium welfare, even if the financial system does not absorb any resources ( $\mu = 0$ ). This is because the presence of a financial sector essentially bids up the price of inputs in terms of liquidity, to the point at which productive entrepreneurs are only able to employ half of the capital stock with their liquidity endowment. The rest of the capital stock is employed by unproductive entrepreneurs, using the financial intermediation technology. Thus, if there is a shock to the financial sector, only half of the capital stock can be employed in production. The economy's reliance on financial intermediation therefore reduces welfare, as financial crises are associated with larger unemployment and lower output.

### 2.3.2 Optimal Policy

I proceed in considering the equilibrium of this economy given a certain tax policy. As in section 2.2.1, I allow for an output tax  $\tau$  and a subsidy  $\epsilon$  on liquidity hoarding. The two tax instruments remain equivalent under this richer environment.

Let  $\tilde{A}$  denote the most profitable use of capital for unproductive entrepreneurs:

$$\tilde{A} = \max\{\underline{A}, (1 - \mu)\bar{A}\} \quad (2.20)$$

And let  $\tau^* = \tau^*(\epsilon)$  and  $\epsilon^* = \epsilon^*(\tau)$  be given by:

$$(1 - \tau^*)\tilde{A}\frac{1}{R^*} = 1 + \epsilon \quad (2.21)$$

$$(1 - \tau)\tilde{A}\frac{1}{R^*} = 1 + \epsilon^* \quad (2.22)$$

Recall that  $R^*$  denotes the price of liquidity under which the productive entrepreneurs have enough liquidity to employ the entire capital stock (equation 2.8).

**Lemma 12** *Proposition 7 generalizes to environments with financial intermediation, with  $\tau^*$  and  $\epsilon^*$  defined as above.*

**Proof:** The economy is equivalent to one in which the distribution of technology is given by:

$$A(x) = \begin{cases} \bar{A} & \text{if } x \in [0, \frac{1}{2}]; \\ \tilde{A} & \text{otherwise.} \end{cases} \quad (2.23)$$

The lemma follows immediately from applying Proposition 7 to this economy.

Consider an economy in which financial intermediation is employed, that is  $\tilde{A} = (1 - \mu)\bar{A}$ . Employing the financial intermediation technology plays the role of self-financing in the environment without intermediation: it is an inefficient use of capital, and should be discouraged. Welfare is improved with a policy that tilts the trade-off in favor of liquidity hoarding. The first best is achieved when all unproductive entrepreneurs hoard liquidity. The financial sector vanishes.

This result may seem counterintuitive in light of the large welfare loss associated with financial crises. In section 2.5.2 I demonstrate that this is because of the partial equilibrium nature of crises: when prices are fixed, a sudden disappearance of financial intermediation is harmful.

Financial intermediation is inefficient in general equilibrium. The presence of liquidity constraints give rise to large profit opportunities; but taking advantage of these opportunities only bids up the price of capital and worsens the liquidity constraints of productive entrepreneurs. By instituting policies which discourage production and favor liquidity hoarding, the price of inputs declines and liquidity constraints are relaxed.

## 2.4 Generality

The principle that encouraging liquidity hoarding (and discouraging production) is welfare-improving in models with liquidity constraints is very general.

Consider the following general setup. Each entrepreneur  $x \in [0, 1]$  is endowed

with a production function,  $F(K, x)$ , where  $F$  is twice differentiable in  $K$  and satisfies  $\frac{\partial F(K, x)}{\partial K} > 0$ ,  $\frac{\partial^2 F(K, x)}{\partial K^2} \leq 0$ , and  $F(K, x)$  is measurable in  $x$ . The liquidity of entrepreneur  $x$  is given by  $Q(x) > 0$ , where  $Q(\cdot)$  is measurable. For completeness, I consider the richer environment in which a financial intermediation technology exists; the parameter  $\mu = 1$  captures the situation in which the returns to intermediation are 0, so the environment is similar to one with no intermediation.

For example, the general setup above can accommodate a model in which all entrepreneurs are endowed with the same decreasing returns technology, but have different amounts of liquidity. This case is interesting as the distributional implications are very different from the one in the two-type example presented in this paper. In the two-type example, the first best allocation is one in which the productive types are the only ones producing. In the model in which all entrepreneurs are endowed with the same decreasing returns technology, the first best allocation has equal production by all entrepreneurs; in both cases, the tax on output (or subsidy on liquidity hoarding) helps bring the liquidity constrained economy closer to the first best allocation.

The main result of this paper generalizes:

**Proposition 8** *Assume that taxes are such that equilibrium output is inferior to the first best. Increasing the tax on output ( $\tau$ ) or the subsidy on liquidity hoarding ( $\epsilon$ ) increases equilibrium output.*

The proof is in the appendix. The line of argument is similar to the particular case of two types.

## 2.5 Realistic Concerns

The unambiguous conclusion of the model in sections 2.2 and 2.3 is that, in an environment with liquidity constraints, a tax on production (or a subsidy on liquidity hoarding) increases output. This seemingly paradoxical conclusion obviously raises some realistic concerns. In this section, I lay out these concerns and discuss their implications for the optimal policy.



### 2.5.1 Elastic Supply of Inputs

A realistic concern with the above proposal is that it depresses the price of inputs. This will adversely affect the incentive to supply inputs, thereby potentially reducing equilibrium output.

Whether the net effect of a production tax is positive or negative depends of course on the elasticity of input supply. If input supply is relatively inelastic, the improvement in resource allocation will dominate. However, if input supply is highly elastic, the decline in input supply may offset the returns from better resource allocation.

Theoretically, this concern can easily be addressed by providing a subsidy for input supply. As output increases with the production tax, in theory there is enough revenue to leave the incentives to supply inputs unchanged, or even to increase them. If a production tax is instituted and a balanced budget is kept, redistributing some of the revenue as an input subsidy will be optimal.

It is worth noting that the proposition to tax production and subsidize input supply emerges naturally in frameworks with liquidity constraints, as it essentially relaxes liquidity constraints by allowing the returns to capital to increase beyond the current supply of liquidity. However, as demonstrated in section 2.2, what matters is not the just the amount of liquidity per-se but rather also the distribution of liquidity among entrepreneurs. An input subsidy cannot, on its own, bring the economy to its first best allocation, as it is unable to change the distribution of liquidity supply in the right way (in fact, a higher capital supply will tend to imply that capital is cheaper from the entrepreneur's perspective, which will encourage more inefficient entrepreneurs to self-finance). In what follows I show that the optimal size of the input subsidy given a specific tax schedule is positive, but only given the optimal policy is it equal to the returns to capital in the competitive equilibrium without liquidity constraints.

I enrich the framework by allowing for a capital supply decision. Assume that capital owners are endowed with a technology to produce capital. The cost of producing  $K$  units of capital is  $e = G(K)$  units of effort, where the cost function  $G(K)$

is convex. The utility of capital owners is modified to include a linear disutility of effort:

$$u(c, e) = c - e \tag{2.24}$$

To summarize, the timeline is modified to include a period  $t = 0$  in which capital owners decide how much capital to supply:

- $t = 0$ : Capital owners produce capital.
- $t = 1$ : Entrepreneurs buy capital from capital owners (and decide whether to self-finance, hoard liquidity, or employ financial intermediation).
- $t = 2$ : Production and taxation. Entrepreneurs consume net-of-tax production and stored liquidity. Capital owners consume the (subsidized) sales revenues from their capital.

**Lemma 13** *Denote by  $\tilde{R}$  the return to the capital owner per unit of supplied capital. In equilibrium, the marginal cost of producing capital is equated with  $\tilde{R}$ :*

$$G'(K) = \tilde{R} \tag{2.25}$$

To see this, note that the capital owners maximize:

$$\max_K \tilde{R}K - G(K) \tag{2.26}$$

The first order condition of the problem above delivers the lemma.

**Corollary 5** *In the absence of a capital subsidy and in the presence of liquidity constraints, capital supply is depressed compared to the first best.*

To see this, note that in the absence of a capital subsidy the return to the capital owner per unit of produced capital is the market price of capital:  $\tilde{R} = R$ . In the first best, the price of capital is  $R^{FB} = \bar{A}$ , whereas in the constrained environment,

$R$  must be lower as a non-trivial set of unproductive entrepreneurs find it optimal to produce. In other words, since not all output can be pledged, the returns to capital are bounded by the liquidity supply.

**Lemma 14** *The supply of capital decreases with the output tax and the hoarding subsidy.*

To prove this lemma, note that the price of capital (as given by equation 2.9) is decreasing in the output tax and the hoarding subsidy, as these decrease the demand for capital. As the price of capital is a component of the compensation of capital (which may also be taxed or subsidized), a lower price implies a lower incentive to supply capital.

This result demonstrates that the optimality of an output tax (or a liquidity subsidy) relies heavily on the assumption that the government is able to subsidize input supply. It has been suggested in the literature that there may be non-trivial commitment problems with the promise to subsidize capital inputs ex-post (as in Kydland and Prescott [1977]): once capital is given, there may be an incentive to expropriate it. In the absence of the ability to subsidize input supply, there is an intermediate optimal output tax, which balances the improvement of resource allocation which the disincentive to supply inputs.

Implicitly, the assumption that the government is able to subsidize capital means that the government is not subject to the same liquidity constraints as the entrepreneurial sector. It is able to pledge post-production tax revenues and credibly commit to pay these as capital subsidies. However, importantly, the fact that the government is able to pledge future revenue is not enough to bring the economy to a first best allocation; it is still necessary to employ a production tax or a hoarding subsidy to overcome the central problem of the distribution of liquidity among entrepreneurs.

It is worth noting that unless the first best policy is implemented (and capital allocation is efficient), the optimal compensation to capital will be less than in the first best. For example, in an equilibrium with no financial intermediation, subsidizing

capital may result in a less efficient allocation of capital ex-post: since the price of capital declines with the capital supply, more entrepreneurs find it optimal to self-finance as the supply of capital increases. The optimal subsidized return to capital ( $\tilde{R}$ ) is therefore less than the average product of capital. While subsidizing capital supply may still be optimal, doing so does not eliminate liquidity constraints, but in some sense even exacerbates them.

### 2.5.2 Sticky Input Prices

In the model presented in sections 2.2 and 2.3, liquidity hoarding is welfare improving; this may seem counterintuitive in light of the recent crisis, in which liquidity hoarding was widely viewed as a source of amplification. While in general equilibrium liquidity hoarding should be encouraged, I demonstrate that in partial equilibrium, when the price of inputs is fixed, liquidity hoarding is harmful. This also suggests that the transition towards a more efficient equilibrium with an output tax or a hoarding subsidy can be costly in the short run. Both points will be demonstrated with the following modification to the model which allows for sticky input prices.

Let  $Y^p$  denote the output produced by capital that was hired with productive entrepreneurs' liquidity, and let  $Y^u$  denote the output produced by capital that was hired with unproductive entrepreneurs' liquidity. Note that total output is:

$$Y = Y^p + Y^u \tag{2.27}$$

**Financial crises.** Assume that the price of inputs,  $R$  is fixed at  $R = R_0$ , that taxes are fixed at  $\tau_0$  and  $\epsilon_0$ , and that  $R_0$  is the equilibrium price given  $\tau_0$  and  $\epsilon_0$ . Assume further that the equilibrium is such that there is financial intermediation in equilibrium. Consider a sudden shock to the intermediation technology, which causes intermediation to be impossible. Alternatively, this could be a “flight to quality” episode, in which savers suddenly acquire a strong preference for unproductive savings over productive investment through the financial sector.

Since financial intermediation is used in equilibrium, it must be the case that

productive entrepreneurs self-finance. Since  $R$  is fixed at  $R_0$ , the amount of capital that productive entrepreneurs can hire with their liquidity endowments remains the same. Therefore, the amount of output produced by capital purchased by productive entrepreneurs remains the same:

$$Y^p = Y_0^p \quad (2.28)$$

However, importantly, the output produced by capital hired by unproductive entrepreneurs declines. This is because unproductive entrepreneurs either switch to liquidity hoarding (which produces no output), or switch to self-financing which is, from their perspective, an inferior production technology. Thus,

$$Y^u < Y_0^u \quad (2.29)$$

It follows that output drops as a result of the financial crisis:

$$Y = Y^p + Y^u < Y_0^s + Y_0^u = Y_0 \quad (2.30)$$

This result suggests that there is an argument for bailing out the financial sector during financial crises, regardless of whether or not the size of the financial sector is efficient in general equilibrium. In partial equilibrium, prices are such that resource allocation is optimal only in the presence of a well-functioning financial sector.

**The short-run implications of an output tax or a hoarding subsidy.** Consider an increase in  $\tau$ ,  $\tau_0 \mapsto \tau > \tau_0$ . According to proposition 7, after prices adjust, the equilibrium is superior. Consider the immediate impact that takes place when  $R$  is fixed.

The implication of the higher production tax is that less entrepreneurs find it optimal to self-finance or engage in intermediation. Once prices adjust, this works towards improving efficiency, because the price of inputs decline and a more selective set of entrepreneurs can each afford to hire more inputs. However, with  $R$  fixed, the

level of production by each of the remaining self-financing entrepreneurs stays the same; as long as prices are fixed, the amount of inputs each entrepreneur can hire is given by his liquidity constraint. The resources previously employed by entrepreneurs who decided to switch to liquidity hoarding become unemployed, and the output produced by them forgone.

Because of price stickiness, an increase in taxes is associated with a drop in output and a rise in unemployment. This suggests that the output tax should be increased gradually, and that the increase should be announced well in advance to allow for the prices of inputs to respond accordingly.

These results trivially apply for a hoarding subsidy as well. An increase in the hoarding subsidy (for example, an increase in the Federal Funds Rate) is recessionary in the short run, while input prices are fixed. However, a sustained high rate of return on unproductive savings increases equilibrium output.

### **2.5.3 The Cost of Intermediation Compared to the Cost of Tax Collection**

The model presented in this paper assumes a certain asymmetry between the private sector and the government: while it is costly for the financial sector to transfer goods from one agent to another, the government is able to collect and redistribute taxes at not cost. In this section I justify this asymmetry.

The assumption that financial intermediation is more costly than tax collection seems empirically relevant. Currently, the amount of resources employed in tax collection is negligible compared to the resources absorbed by the financial sector. The recent financial crisis as well as other occurrences of financial fragility suggest that the social cost of financial intermediation may in fact be even larger than that calculated based on employment and profit shares.

Realistically, instituting a large tax on production may increase tax collection costs substantially, as the incentives to evade taxes are higher. In this section I present evidence that suggests that, at least for a small increase in the production tax, there

will be no increase associated with collection costs. Since, by Lemma 12, output is *monotonically* increasing in the production tax, this suggests a benefit to increasing the production tax at least by a small amount. Further, I suggest that the marginal savings generated from reducing the size of the financial sector are potentially large.

Though it sounds plausible, there is no clear evidence that the direct cost of tax collection increases with tax revenues. Yesin [2004] provides data from the IRS Data Book regarding the operating costs and total tax collection for the years 1976-2000 (table 2 in Yesin [2004]). Using these figures, as well as US nominal GDP from the FRED database, I calculate the operating costs and the tax revenues as a fraction of GDP. Figure 2-1 presents the linear regression of the share of operating costs on the share of collection. If anything, this relationship is decreasing.

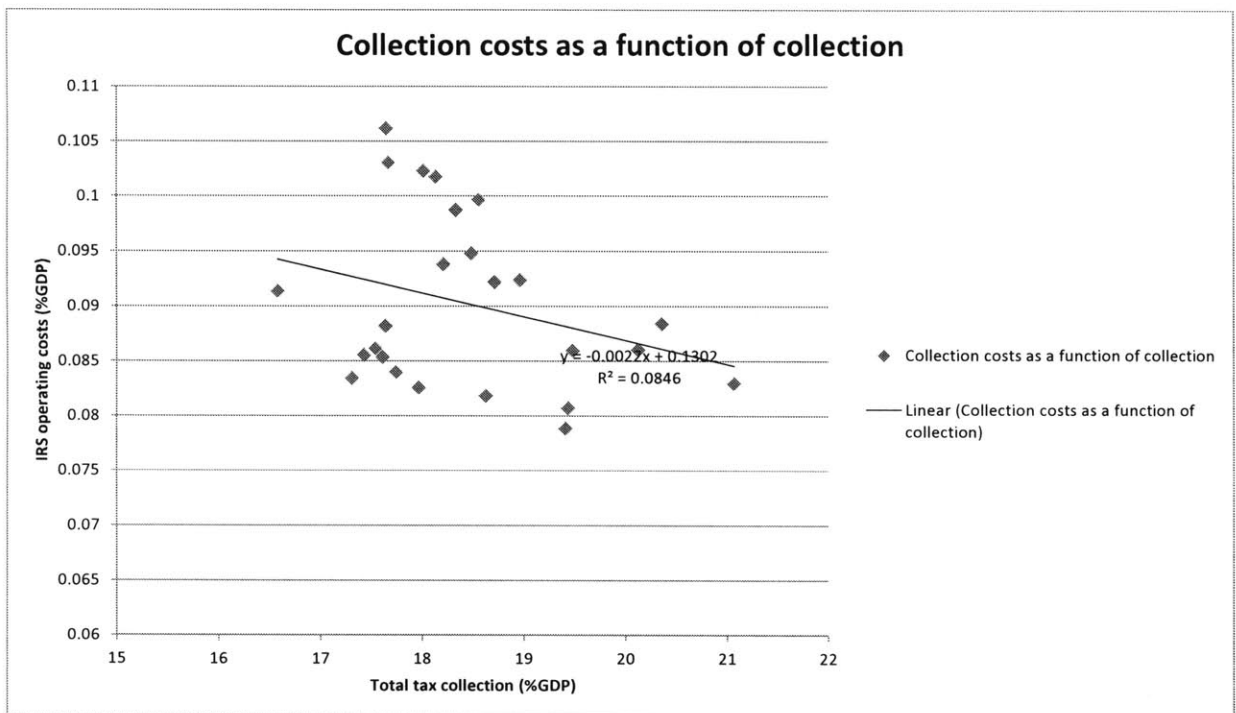


Figure 2-1: Tax costs

Similarly, a small cross-country comparison suggests that the operating costs of the tax collection agency do not increase dramatically with the average tax rate.

Table 3 in Yesin [2004] provides figures for the total tax revenues and the budget of the tax collection agency for a few countries. Using WEO data, I calculate the total tax revenues and the budget of the tax collection agency as a percent of GDP. The results in in table 2.4.<sup>8</sup> Regressing the budget of the collection agency on the tax revenue results in a slightly negative coefficient. Indeed, the numbers suggests that the average tax rate is not the primary determinant of the collection costs. For example, Australia spends 0.7% of GDP collecting an average tax rate of 24%. Norway spends only 0.3% of GDP collecting an average tax rate of 40%.

Table 2.4: Tax revenues and the budgets of collection agencies

Country	Total tax revenue (% GDP)	Budget of tax collection agency (%GDP)
Australia	24	0.7
Canada	16	0.4
Israel	27	0.3
Norway	40	0.3
US	19	0.1

Of course, to determine how collection costs would change with an increase in taxes requires more rigorous econometric analysis. However, both the time series comparison for the US and the cross country comparison suggest that the marginal cost of additional taxation is not high.

Moreover, it is not necessarily the case that in order to increase the production tax the total tax revenue must increase. Rather, the production tax can (and in the framework of this model, should) substitute for various taxes on inputs, such as labor income taxation and capital taxation (which, in this model, should be subsidized). A better resource allocation can be achieved simply by changing the composition of taxes, without changing the total tax revenue.

I now turn to the marginal cost of financial intermediation. According to Friedman [2009], from the 1950s to the 1980s, the finance sector accounted for 10% of all profits earned by US corporations. According to Friedman, this share has risen to 35% in the years 2000-2005. This provides some measure of the increase in the cost of

<sup>8</sup>I exclude Turkey, because the figures reported in Yesin [2004] combined with the current priced GDP provided by the WEO produce unreasonable figures of tax collection way above 100% of GDP.



financial intermediation. A crude measure of the increase in the quantity of financial intermediation is the increase in the bank credit of commercial banks. Using data from the FRED database<sup>9</sup>, I estimate that bank credit as a percent of GDP increased by about 160% from the 1970's to the years 2000-2005. Thus, an increase of 160% in intermediation is associated with a 250% increase in costs. This suggests that increasing the quantity of intermediation by 1% increases the costs of operating the financial sector by about 1.5%. The positive marginal cost of financial intermediation implies that downsizing the extent of financial intermediation will reduce its costs substantially.

Further, note that the figures above reflect only the normal operating costs of the financial sector, and do not include the costs of the fragility induced by financial crises. It is likely that with a smaller financial sector, the vulnerability of the economy to financial crises will be reduced. Moreover, it is possible to argue that the crisis itself was due to excess capacity in the financial sector - in which case, the marginal benefit of reducing the size of the financial sector is very high.

To summarize, since the marginal increase in collection cost is negligible compared to the marginal savings generated from crowding out the financial sector, the net effect of a marginal increase in the production tax is positive (note that the crowding out of the financial sector is a lower bound on the size of the benefits from increasing the production tax, as the production tax may also imply a better equilibrium allocation of capital among entrepreneurs):

$$\frac{\partial Y}{\partial \tau} \geq \underbrace{\frac{\partial \text{Collection costs}}{\partial \text{tax revenues}}}_{=0} \underbrace{\frac{\partial \text{tax revenues}}{\partial \tau}}_{>0} - \underbrace{\frac{\partial \text{Financial sector}}{\partial \text{Intermediation}}}_{>0} \underbrace{\frac{\partial \text{Intermediation}}{\partial \tau}}_{<0} > 0 \quad (2.31)$$

---

<sup>9</sup>Series: TOTBKCR, Bank Credit of All Commercial Banks

## 2.6 Conclusion

The prevailing sense that the size of the financial sector is “too large” may at first seem at odds with traditional economic theory. Typically, in hyper-competitive environments such as the financial sector, the profits of the sector reflect its contribution to output. However, as Paul Volcker elegantly expresses in *The New York Review of Books*, it is questionable whether the financial sector indeed adds surplus which can justify its profits:

“Has the contribution of the modern world of finance to economic growth become so critical as to support remuneration to its participants beyond any earlier experience and expectations? Does the past profitability of and the value added by the financial industry really now justify profits amounting to as much as 35 to 40 percent of all profits by all US corporations? Can the truly enormous rise in the use of derivatives, complicated options, and highly structured financial instruments really have made a parallel contribution to economic efficiency?”

-Paul Volcker, ‘The Time We Have Is Growing Short’, *The New York Review of Books*, June 24th 2010.

This paper suggests that the answer to the questions above is negative. The profits of the financial sector do not reflect its contribution to output. Rather, people are willing to pay for intermediation not because of the surplus it generates, but because it transfers surplus to people with wealth who do not have highly productive investment opportunities of their own. The payment to the financial sector reflects the return to intermediation from the perspective of its clients. An immediate conclusion of this is that if the government can freely institute transfers, intermediation is wasteful as the transfer of wealth can be done more cheaply through taxation. Realistically, even if the government is unable to transfer resources freely between individuals, it is likely that the transfer implied by financial intermediation is a socially undesirable one, as it transfers surplus from liquidity constrained entrepreneurs and small businesses to wealthy liquidity suppliers.

Unlike the simplified model presented here, in reality there are limited high-profit investment opportunities. This further increases the amount spent on intermediation, as resources must be spent in order to compete for the opportunity to take advantage of these. Of course, this rat-race is wasteful from a social perspective. As Friedman [2009] points out:

“Perversely, the largest individual returns seem to flow to those whose job is to ensure that microscopically small deviations from observable regularities in asset price relationships persist for only one millisecond instead of three. These talented and energetic young citizens could surely be doing something more useful.”

-Benjamin Friedman, ‘Overmighty finance levies a tithe on growth’. *The Financial Times*, August 26 2009.

The inefficiency generated by the large amount of resources spent on financial intermediation calls for structural change. The model in this paper suggests that a simple tax policy can help reduce the size of the financial sector, while improving the equilibrium allocation of capital. Previous policy proposals have focused on the role of a transaction tax, which essentially taxes intermediation. In this paper it is demonstrated that while this may improve welfare in some equilibria, it may also be harmful as it encourages inefficient self-financing. The unambiguous recommendation that emerges from this paper is a simple tax on production, coupled with a subsidy on unproductive savings (“liquidity hoarding”) and a subsidy on input supply.

A tax on production and a subsidy on unproductive savings discourage inefficient usage of capital. Only entrepreneurs with truly productive investment opportunities will find it optimal to self-finance; similarly, only intermediation which is relatively cheap and highly productive will still be profitable in equilibrium. This improves the equilibrium allocation of capital among entrepreneurs, and increases the share of capital employed in the productive sector. The subsidy on input supply helps correct the incentive to supply capital when the returns to capital are depressed.

In addition to the benefits mentioned above, an economy that relies less on financial intermediation is, at the very least, less exposed to financial crises; possibly, financial crises are eliminated entirely as intermediation becomes simpler and less profitable. The possibility of using simple policies to reduce the need for financial intermediation without compromising on the welfare improving resource allocation associated with the financial sector is particularly appealing given that the real cost of financial crises is large.

# Bibliography

- S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, August 1994.
- S. Rao Aiyagari. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy*, 103(6):1158–1175, December 1995.
- S. Rao Aiyagari and Ellen R. McGrattan. The optimum quantity of debt. *Journal of Monetary Economics*, 42:447–469, 1998.
- George J. Benston and Clifford W. Smith. A transactions cost approach to the theory of financial intermediation. *Journal of Finance*, 31(2):215–231, May 1976.
- John Boyd and Edward Prescott. Financial intermediary coalitions. *Journal of Economic Theory*, 38:211–232, 1986.
- Ricardo J. Caballero and Arvind Krishnamurthy. International and domestic collateral constraints in a model of emerging market crises. *Journal of Monetary Economics*, 48:513–548, 2001.
- Ricardo J. Caballero and Arvind Krishnamurthy. Collective risk management in a flight to quality episode. *Journal of Finance*, 63(5):2195–2230, October, 2008.
- Ricardo J. Caballero and Alp Simsek. Fire sales in a model of complexity. 2010.
- Tim S. Campbell and William A. Kracaw. Information production, market signalling, and the theory of financial intermediation. *Journal of Finance*, 35(4):863–882, September 1980.

- Julio Davila, Jay H. Hong, and Per Krusell Jose-Victor Rios-Rull. Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks. 2007.
- Douglas Diamond. Financial intermediation and delegated monitoring. *Review of Economic Studies*, 51:393–414, 1984.
- Douglas W. Diamond and Philip H. Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419, June 1983.
- Benjamin Eden. The role of government in the consumer loans market. 2010.
- Emmanuel Farhi, Mikhail Golosov, and Aleh Tsyvinski. A theory of liquidity and regulation of financial intermediation. *Review of Economic Studies*, 76:973–992, 2009.
- Benjamin Friedman. Overmighty finance levies a tithe on growth. *The Financial Times*, August 26 2009.
- Milton Friedman. The optimum quantity of money. In *The Optimum Quantity of Money and Other Essays*, pages 1–50. Chicago: Aldine, 1969.
- Gary Gorton and Lixin Huang. Liquidity, efficiency, and bank bailouts. *American Economic Review*, 94(3):455–483, June 2004.
- Gary Gorton and Andrew Winton. Financial intermediation. In George Constantinides, Milt Harris, and Rene Stulz, editors, *Handbook of the Economics of Finance*, volume 1, chapter 8, pages 431–552. Elsevier, 2003.
- Bengt Holmström and Jean Tirole. Private and public supply of liquidity. *Journal of Political Economy*, 106(1):1–40, February 1998.
- Bengt Holmström and Jean Tirole. Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics*, 112(3):663–691, August 1997.

- Charles J. Jacklin. Demand deposits, trading restrictions, and risk sharing. In Edward C. Prescott and Neil Wallace, editors, *Contractual Arrangements for Intertemporal Trade*, pages 26–47. Minneapolis: University of Minnesota Press, 1987.
- Timothy J. Kehoe and David K. Levine. Debt-constrained asset markets. *Review of Economic Studies*, 60(4):865–888, October 1993.
- Nobuhiro Kiyotaki and John Moore. Credit cycles. *Journal of Political Economy*, 105(2):211–248, 1997.
- Finn E. Kydland and Edward C. Prescott. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85(3):473–492, June 1977.
- Hayne E. Leland and David H. Pyle. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance*, 32(2):371–387, May 1977.
- Florian Scheuer. Pareto-optimal taxation with aggregate uncertainty and financial markets. 2009.
- Robert Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21:265–293, 1979.
- Paul Volcker. 'the time we have is growing short'. *The New York Review of Books*, LVII(11), June 24 2010.
- Stephen Williamson. Costly monitoring, financial intermediation, and equilibrium credit rationing. *Journal of Monetary Economics*, 18:159–179, 1986.
- A. Pinar Yesin. Tax collection costs, tax evasion and optimal interest rates. 2004.

## 2.7 Appendix

### 2.7.1 Proof of Lemma 10

First, note that the price of capital,  $R$ , is given by the following equation:

$$Q(1 - h) = R\bar{K} \quad (2.32)$$

Where  $h$  is the fraction of hoarded liquidity. As productive entrepreneurs always strictly prefer to produce,  $h \leq \frac{1}{2}$ . It follows that:

$$(1 - h) \geq \frac{1}{2} \quad (2.33)$$

Hence,

$$R = \frac{(1 - h)Q}{\bar{K}} \geq \frac{\frac{1}{2}Q}{\bar{K}} \quad (2.34)$$

The unproductive entrepreneur's return to production is  $\underline{A}\frac{1}{R}$ . The return to liquidity hoarding is 1. Hence, entrepreneurs will choose to self-finance if:

$$\underline{A}\frac{1}{R} > 1 \Leftrightarrow \underline{A} > R \geq \frac{\frac{1}{2}Q}{\bar{K}} \quad (2.35)$$

### 2.7.2 Proof of Lemma 11

I first prove that only productive entrepreneurs use intermediated capital, given the assumption that the repayment rate is  $r = \bar{A}$ . Assume by way of contradiction that there is an unproductive entrepreneur that uses intermediated capital. Thus, there is an agent  $y$  who finds it optimal to deposit capital in the hands of an unproductive entrepreneur. This means that the repayment agreement is at least  $r = \bar{A}$ , because otherwise  $y$  would do better by depositing the capital in the hands of a productive entrepreneur. But, as the return to capital in the hands of the unproductive entrepreneur is only  $\underline{A} < \bar{A}$ , it follows that the unproductive entrepreneur loses from this contract and would not agree to this repayment rate.

Next, I prove that the repayment rate is  $r = \bar{A}$ . For  $r > \bar{A}$ , productive en-



entrepreneurs would not agree to the contract. Assume by way of contradiction that  $r < \bar{A}$ . The demand for intermediated capital is strictly positive, as productive entrepreneurs profit from intermediated capital. It follows that  $r = \bar{A}$ .

### 2.7.3 Proof of Proposition 8

Consider an equilibrium in which the output is less than first best. For simplicity assume that the first best allocation is unique. The fact that output is inferior to the first best implies that there is a set of entrepreneurs of positive measure,  $E$ , such that these entrepreneurs employ less capital than they do in the first best.

It follows that for  $x \in E$ :

$$\frac{\partial F(K, x)}{\partial K} > R^{FB} > R \quad (2.36)$$

The liquidity constraint is binding for these entrepreneurs, so they use all of their liquidity to self-finance. For any  $x \in E$ , the capital self-employed by entrepreneur  $x$  is:

$$Q(x) = RK(x) \quad (2.37)$$

Note that  $\tau$  and  $\epsilon$  do not change the tradeoff between self-financing and intermediation, but only the return to liquidity hoarding relative to these two options.

Denote by  $r$  the return to intermediation.

**Claim 3** *For every  $x \in E$ ,*

$$(1 - \mu)r < \frac{\partial F(K_x, x)}{\partial K} \quad (2.38)$$

*And:*

$$r = \max_{x \in E} \frac{\partial F(K_x, x)}{\partial K} \quad (2.39)$$

*Where  $K_x$  is the total capital employed by the production technology of entrepreneur  $x$ .*

The first inequality stems from the optimality of entrepreneur  $x$ 's decision to

self-finance. The second equality results from the optimal decision of where to intermediate capital.

The demand for capital from  $x \in E^c$  (where  $E^c$  denotes the complement of  $E$ ) is given by:

$$\max_{K(x), K_i(x)} (1 - \tau)(F(K(x), x) + r(1 - \mu)K_i(x)) + (1 + \epsilon)(Q(x) - (K(x) + K_i(x))) \quad (2.40)$$

s.t.:

$$R(K(x) + K_i(x)) \leq Q(x) \quad (2.41)$$

The demand for capital ( $K(x) + K_i(x)$ ) decreases with  $\tau$  and  $\epsilon$ . It follows that the equilibrium price of capital is lower when  $\tau$  and  $\epsilon$  are higher, so that entrepreneurs in  $E$  can hire more capital. Output increases.

# Chapter 3

## Global Imbalances in a Monetary Model of Liquidity Constraints

### Abstract

I construct a monetary model in which liquidity constraints emerge in equilibrium. For fast-growing economies, financial integration is associated with lower consumption and increased consumption volatility. Foreigners extract rents from supplying liquidity to the constrained productive sector. Input prices appreciate, but equilibrium output remains unchanged. This results in lower consumption, as some of the output is used as payment to liquidity suppliers. The magnitude of the flows implied by the model are roughly 2%-6% of GDP. In the presence of sticky input prices, the reliance on foreign liquidity supply is a potential source of instability, as contractions in foreign liquidity supply lead to drops in employment, output and consumption.

**JEL Classification:** E40, E50, F30, G15

**Keywords:** global imbalances, cash in advance, liquidity constraints, financial integration, instability, exchange rate manipulation

### 3.1 Introduction

When firms are liquidity constrained, there are opportunities for liquidity suppliers to extract rents. In this paper I explore the implications of this in a global equilibrium context, in which some fast-growing regions are liquidity constrained while other regions are at their steady state and are able to supply liquidity. The model predicts

an increasing flow of goods from the constrained region to the unconstrained region, consistent with “global imbalances”. From a normative perspective, this model suggests that financial integration is welfare-reducing for the constrained regions, but beneficial for the unconstrained regions.

I consider a monetary model in which there is a cash in advance constraint on employing production inputs, which I refer to as a liquidity constraint. In the closed economy, a binding liquidity constraint offers a positive rate of return on holding cash, guaranteeing that firms will be willing to carry cash from one period to the next. A feature of this model is that faster-growing economies will be more constrained in equilibrium, in the sense that the marginal product of inputs is higher relative to their price.

I study the effects of financial integration between a fast growing constrained economy and a steady state rest of the world. Financial integration is assumed to be limited in the sense that only short term borrowing and lending from abroad is permitted: firms can borrow within a period to bridge the gap between production expenses and sales, but households cannot borrow across periods to smooth consumption. In the integrated environment, the returns to holding cash are pinned down by the rest of the world at a relatively low level. Given their increasing consumption path, domestic shareholders are no longer willing to hold cash at the low rate of return. Thus, in equilibrium, the financing of production expenses is done entirely through foreign liquidity. The rents paid to foreign liquidity suppliers reduce equilibrium consumption compared to the closed economy equilibrium. Further, the economy’s production becomes entirely reliant on foreign liquidity supply, which may be source of instability.

The mechanics of this model are close to Eden [2011b]. In Eden [2011b], I consider a closed economy with liquidity constraints in which the price of production inputs in terms of liquidity is determined in equilibrium. The equilibrium determination of this price is the key ingredient in both models. There are two main differences between this paper and Eden [2011b]: first, the focus of the two papers is different. Here, I focus on the welfare implications of financial integration and on global imbalances.

In Eden [2011b] I focus on the general equilibrium costs of financial intermediation in a closed economy context. Second, this paper extends Eden [2011b] by endogenizing the liquidity supply in the context of a dynamic, monetary model. In Eden [2011b], I consider a single-period model in which liquidity is given exogenously. Here, I demonstrate how a binding liquidity constraint can arise in equilibrium, and how it changes following financial integration.

This model implies an increasing flow of goods from constrained economies to the liquidity supplying regions. This model of “global imbalances” is most closely related to the equilibrium view in Caballero et al. [2008], in which financial frictions lead to a permanent flow of savings towards economies with better savings facilities. While the essence of the financial friction is similar - inability to pledge future output - the implications are quite distinct. In Caballero et al. [2008], the driving force behind the global imbalances is the demand for savings from emerging economies. In contrast, in this model, while the central bank ends up accumulating foreign reserves, there is no demand for savings. In fact, if they had access to long term international borrowing and lending, domestic agents would choose to borrow to smooth consumption. The accumulation of foreign reserves is just an artifact of the increasing demand for converting foreign into domestic currency in order to be able to supply liquidity to domestic producers.

There are other papers in which agents face endogenous borrowing constraints, such as Holmstrom and Tirole [1997] and, more recently in the context of financial integration, Korinek [2010]. In these papers, borrowing constraints result from a moral hazard problem and incomplete enforcement. Here, while it is assumed that future output is not pledgable and inputs must be purchased with cash, enforcement does not play an important role. Rather, the binding cash in advance constraint is necessary in order for people to be willing to hold cash across periods. In contrast to this literature, a binding liquidity constraint does not necessarily imply a loss of welfare. This paper provides an alternative model for binding liquidity constraints, that is immune to the critique in Aiyagari [1994] that, in a dynamic model, liquidity constraints vanish in the long run as firms accumulate capital.

This paper is also related to the literature on cash in advance constraints in the productive sector, such as Stockman [1981] or Abel [1985]. Most closely related to this paper is Neumeyer and Perri [2005], who consider an open economy with a “working capital” constraint which requires that part of the wage bill must be financed with foreign credit. This paper endogenizes this restriction in a model in which there is a cash in advance constraint on hiring labor. In equilibrium firms must rely on foreign liquidity suppliers for cash.

Finally, this paper is broadly related to the literature on the distributional implications of opening to trade and the welfare implications of liberalization. Bhagwati and Brecher [1980] present the idea that in the presence of foreign-owned inputs of production, trade liberalization can decrease domestic welfare if it increases the relative wage of the inputs that are disproportionately owned by foreigners. In this model, the distributional implications of liberalization go in the opposite direction: financial integration increases the return to domestic labor, while it lowers the return to liquidity that can be supplied by foreigners. However, welfare decreases as the economy imports liquidity in equilibrium, despite the fact that the economy’s liquidity needs can be satisfied in autarky at no aggregate cost. The distributional implications of financial integration are in the spirit of Antras and Caballero [2009]: the returns to liquidity decline, as liquidity can be imported more cheaply from abroad, allowing labor to absorb a higher share of output. However, in contrast to Antras and Caballero [2009], the net surplus of the economy declines. The key difference is that liquidity is not a real factor of production, in the sense that importing liquidity does not allow the economy to expand its production.

The rest of the paper is organized as follows. In section 3.2, I illustrate the main insights using a single period real model, where liquidity constraints are assumed exogenous. In section 3.3, I endogenize the liquidity constraints in a dynamic monetary model. In section 3.4 I solve for the closed economy equilibrium. In section 3.5 I consider the implication of financial integration. I show that financial integration is associated with lower consumption. In section 3.6 I explore the implications for the current account. In section 3.7 I show that the integrated equilibrium is one in

which the domestic economy is vulnerable to shocks to foreigner's willingness to supply liquidity, and demonstrate that the central bank has an incentive to depreciate the currency in such an event. In section 3.8 I conclude.

## 3.2 Single Period Real Model

I begin by illustrating the main insights using a simplified single period real model. There is a unit measure of households and a unit measure of firms. Each household supplies  $L$  units of labor inelastically. Each household owns a share of the productive sector. The technology is given by  $\rho F(L^e)$ , where  $L^e$  is employed labor and  $F(\cdot)$  is an increasing function with diminishing returns ( $F'(\cdot) > 0$ ,  $F''(\cdot) < 0$ ).

At the beginning of the period, each firm is endowed with  $Q$  units of current output which will be referred to as *liquidity*. Firms can trade in liquidity. The amount of liquidity demanded by firm  $i$  (in addition to its endowment) is denoted  $Q_i^d$ , where  $Q_i^d$  can be positive or negative. The total amount of liquidity held by the firm at the beginning of the period is therefore  $Q + Q_i^d$ .

The price of liquidity is set in terms of output at the end of the period: one unit of liquidity at the beginning of the period costs  $1 + r$  units of the final good, to be delivered after production takes place. Under autarky,  $r^{aut}$  is set such that the market for liquidity clears; in the open economy, it is assumed that  $r$  is exogenously fixed by the rate of return to liquidity in global markets.

Denote by  $L_i^e$  the amount of labor employed by firm  $i$ , and let  $w$  be the equilibrium wage.

### 3.2.1 The Unconstrained Benchmark

Before adding the liquidity constraint, it is useful to consider the benchmark of the unconstrained economy. In the unconstrained economy, the firm's problem is:

$$\max_{L_i^e, Q_i^d} \rho F(L_i^e) + Q + Q_i^d - (1 + r^{aut})Q_i^d - wL_i^e \quad (3.1)$$

In equilibrium, firms choose labor such that the marginal product of labor is equal to the wage. By symmetry, labor market clearing implies that  $L_i^e = L$  for all  $i$ . The wage is therefore given by:

$$w = \rho F'(L_i^e) = \rho F'(L) \quad (3.2)$$

Under autarky, liquidity market clearing implies that  $r^{aut} = 0$ : otherwise, if  $r^{aut} < 0$ , firms demand an infinite amount of liquidity; if  $r^{aut} > 0$  firms want to supply an infinite amount of liquidity. Having  $r^{aut} = 0$  guarantees that each firm demands 0 additional units of liquidity ( $Q_i^d = 0$ ).

As the entire labor force is employed in equilibrium, and the marginal product of labor is equated across firms, output is equal to  $\rho F(L)$ . Output is distributed both as wages and as dividends and is consumed in its entirety.

Dividends are given by:

$$d = \rho F(L) + Q - wL \quad (3.3)$$

The first best level of consumption is therefore given by:

$$c^{FB} = wL + d = wL + \rho F(L) + Q - wL = \rho F(L) + Q \quad (3.4)$$

Note that the unconstrained economy necessarily gains from financial integration. To see this, note that the derivative of the firm's problem with respect to  $Q_i^d$  is  $-r$ . If  $r > 0$ , firms would export liquidity at an arbitrarily large amount, resulting in large profits. If  $r < 0$ , firms would import liquidity at an arbitrarily large amount. Profits would be arbitrarily large in either case. If  $r = 0$ , prices remain at their autarkic level; the economy therefore (weakly) gains from financial integration, regardless of the international rate of return to liquidity.



### 3.2.2 The Constrained Economy

I continue by describing the equilibrium in the constrained economy, in which labor inputs must be purchased in advance of production.

Firms must purchase labor with liquidity; the amount of labor employed by firm  $i$  must satisfy:

$$wL_i^e \leq Q + Q_i^d \quad (3.5)$$

Firms therefore face the following optimization problem:

$$\max_{L_i^e, Q_i^d} \rho F(L_i^e) + Q + Q_i^d - (1+r)Q_i^d - wL_i^e \quad (3.6)$$

s.t.

$$wL_i^e \leq Q + Q_i^d \quad (3.7)$$

**The equilibrium under autarky.** It turns out that for  $\rho$  sufficiently large, the liquidity constraint (equation 3.7) is binding in equilibrium. The following lemma will be useful in deriving this result:

**Lemma 15** *Under autarky, the liquidity constraint (equation 3.7) is binding if and only if  $r^{aut} > 0$ . If the liquidity constraint is not binding,  $r^{aut} = 0$ .*

The proof of this lemma, together with other omitted proofs, is in the appendix.

**Claim 4** *For  $\rho$  sufficiently large, the liquidity constraint is binding in equilibrium.*

To prove this claim, I will show that a 0 rate of return on liquidity ( $r^{aut} = 0$ ) will generate excess demand for liquidity. As, by Lemma 15,  $r^{aut} = 0$  is the only rate of return consistent with an unconstrained equilibrium, it will follow that the liquidity constraint must bind in equilibrium.

Consider the firm's problem above, holding the wage  $w$  constant, and assuming that  $r^{aut} = 0$ . Ignoring the liquidity constraint, the firm would like to set the marginal product of labor equal to the wage. Note that the marginal product of labor is

increasing in  $\rho$  - thus, the higher  $\rho$ , the more labor the firm wants to hire. For  $\rho$  sufficiently large the liquidity constraint is binding for  $Q_i^d = 0$ , that is:

$$\rho F'(L_i^e = \frac{Q}{w}) \geq w \quad (3.8)$$

The amount of labor  $L_i^e = \frac{Q}{w}$  is the maximum amount of labor that the firm can hire given its liquidity endowment. Thus, for any *given*  $w$ , for a sufficiently high  $\rho$  the rate of return  $r^{aut} = 0$  generates excess demand for liquidity. Thus, in equilibrium, it must be the case that  $r^{aut} > 0$  and that the liquidity constraint is binding.

Consider next the equilibrium determination of the wage  $w$ . Whenever the firm's liquidity constraint is binding, the wage  $w$  is determined by the aggregate liquidity constraint, that is:

$$wL = Q \Rightarrow w = \frac{Q}{L} \quad (3.9)$$

From the analysis of the firm's problem, for any given  $w$  the liquidity constraint is binding for a sufficiently large  $\rho$ . Thus, for  $\rho$  sufficiently large, the liquidity constraint binds for  $w = \frac{Q}{L}$ . In this case, the liquidity constraint is binding in equilibrium, as the wage is determined by the aggregate liquidity constraint and - given that wage - all firms use their entire liquidity endowment to purchase labor.

The price of liquidity is determined in equilibrium to assure that the market for liquidity clears. The price of liquidity is equal to the additional revenue that it generates. An additional unit of liquidity can buy  $\frac{1}{w}$  units of labor; each unit of labor produces  $\rho F'(L)$  units of output. The price of liquidity is therefore given by:

$$1 + r^{aut} = \frac{\rho F'(L)}{w^{aut}} = \frac{\rho F'(L)L}{Q} \quad (3.10)$$

It turns out that despite the fact that the liquidity constraint is binding, output is still at its first best level. To see this, consider the problem of a planner trying to maximize output:

$$Y^{FB} = \max_{L_i^e} \int_0^1 \rho F(L_i^e) di \quad (3.11)$$

s.t.

$$\int_0^1 L_i^e di \leq L \quad (3.12)$$

Given the assumption that  $F(\cdot)$  is increasing and concave, the solution to this problem is that output is maximized when the economy is at full employment and the marginal product of labor is equated across firms.

These conditions are satisfied in the constrained economy. As labor is supplied inelastically, the entire labor force is employed in equilibrium. The fact that there is a market for liquidity guarantees that the marginal product of labor is equated across firms: all firms equate the return to liquidity with its price<sup>1</sup>:

$$\frac{F'(L_i^e)}{w} = 1 + r^{aut} \Rightarrow F'(L_i^e) = F'(L_j^e) \quad (3.13)$$

In this model, output is at its first best level, and consumption is therefore also at its first best level. To see this, denote dividend income by  $d$ . In the autarkic equilibrium,  $d$  is given by aggregate output plus the initial liquidity endowment, net of the wage bill:

$$d = \rho F(L) + Q - wL \quad (3.14)$$

Consumption is given by:

$$c^{aut} = wL + d = wL + (\rho F(L) + Q - wL) = \rho F(L) + Q = c^{FB} \quad (3.15)$$

To conclude, under autarky the presence of the liquidity constraint has no effect on welfare. Consumption is at its first best level, as inflated dividend income exactly offsets the effect of depressed labor income.

---

<sup>1</sup>The assumption that firms have identical technologies and identical liquidity endowments is not important for this result. It is easy to show that the result generalizes to settings in which the liquidity endowments  $Q_i$  are bounded away from 0 ( $Q_i > \epsilon > 0$  for all  $i$ ) and  $F_i(\cdot)$  changes continuously across  $i$  in the bounded set  $i \in [0, 1]$ .

**Financial integration.** In contrast to the unconstrained case, when the economy is constrained financial integration can decrease equilibrium welfare. Consider the implications of having the constrained economy open to international liquidity flows. Assume that the return to liquidity on global markets is pinned down by  $1 + r$ , where  $1 + r$  is small compared to the return to liquidity in the closed economy but still greater than 1<sup>2</sup>:

$$\frac{\rho F'(L)}{w^{aut}} = 1 + r^{aut} > 1 + r > 1 \quad (3.16)$$

**Claim 5** *Financial integration reduces domestic consumption.*

As long as  $\frac{\rho F'(L)}{w} > 1 + r$ , there are rents from supplying liquidity to constrained firms. Thus, the wage appreciates until  $\frac{\rho F'(L)}{w} = 1 + r$ . The higher wage bill is financed by foreign liquidity supply. Denote the equilibrium supply of foreign liquidity by  $Q^*$ . The wage bill now satisfies:

$$wL = Q + Q^* \quad (3.17)$$

Similarly to the autarkic equilibrium, output is still at its first best level. However, welfare declines as equilibrium consumption is lower. This is because some of the output has to be paid as rents to foreign liquidity suppliers:

$$d = \rho F(L) + Q + Q^* - wL - (1 + r)Q^* \quad (3.18)$$

$$c = wL + d = wL + (\rho F(L) + Q + Q^* - wL - (1 + r)Q^*) \quad (3.19)$$

$$= \rho F(L) + Q - rQ^* < c^{aut} \quad (3.20)$$

Financial integration reduces equilibrium consumption because constrained firms fail to internalize the effect of their borrowing on the equilibrium wage. From the perspective of each firm faced with a tight liquidity constraint, the prospect of borrowing cheaply from abroad to finance the purchase of more labor is an attractive one. However, as all labor is already employed, this activity only bids up the price

---

<sup>2</sup>When this condition is violated, the constrained economy will gain from financial integration. However, in section 3.3 it will be shown that this condition will be satisfied in equilibrium for relatively fast-growing economies.

of labor. As the wage bill becomes unaffordable to the domestic productive sector, some of the returns to labor must be paid as rents to foreign liquidity suppliers.

This model also admits to the following features:

- **Global imbalances.** Note that this model predicts a net flow of goods from constrained economies to foreign liquidity suppliers of the amount of  $(1+r)Q^* - Q^* = rQ^*$ . This “global imbalance” is a result of equilibrium rents from supplying liquidity to constrained economies.

From the government’s perspective, there is an incentive to issue new liquidity to crowd out foreign liquidity supply. In the monetary economy, it will be shown that if the central bank can issue currency that is not backed by foreign reserves, it can improve welfare by distributing liquidity directly to firms. However, this policy will not be sufficient to eliminate global imbalances: the rents paid to foreigners will still be positive even under the optimal policy.

- **Instability associated with financial integration.** In the integrated equilibrium, the productive sector relies on foreign liquidity supply in order to finance the wage bill. This implies that, if wages are sticky, a sudden reluctance of foreigners to supply liquidity will induce a drop employment. To illustrate this, assume that the interest rate at which foreigners are willing to lend jumps to  $r' > r$ , and that  $w = \hat{w}$  is constant at its “no shock” equilibrium level. Employment must drop so that  $L^e$  satisfies  $\frac{\rho F'(L^e)}{\hat{w}} = 1 + r'$ . Drops in output and consumption will follow as well.

Given this shock, the government has an incentive to increase employment by depreciating the currency, thereby making domestic labor effectively cheaper to foreigners. The price of one unit of labor in terms of foreign currency is  $\frac{\hat{w}}{e}$ , where  $e$  is the exchange rate. The domestic rate of return is therefore  $\frac{\rho F'(L^e)e}{\hat{w}} = 1 + r'$ . By increasing  $e$ , the government can induce higher employment.

### 3.3 A Dynamic Monetary Model of Liquidity Constraints

The key novel ingredient in this model is the equilibrium determination of the price of production inputs in terms of liquidity. The relevance of this mechanism therefore depends on the equilibrium determination of liquidity and how it responds to changes in the environment. In this section I construct a monetary model in which liquidity is the amount of money held by firms at the production stage. Both in the closed economy and in the integrated equilibrium, binding liquidity constraints are a necessary feature of the environment.

I consider a discrete time infinite horizon model, where time periods are indexed by  $t = 1, 2, \dots$

**Labor supply.** There is a unit measure of households that supply  $L$  units of labor inelastically each period. I assume for simplicity that labor is the only input of production.

**Technology.** The production technology in period 0 is given by  $F_0(L)$ , where  $F'_0 > 0$  and  $F''_0 < 0$ . I assume constant productivity growth, governed by  $\rho > 1$ :

$$F_t(\cdot) = \rho^t F_0(\cdot) \tag{3.21}$$

**Productive sector.** There is a competitive productive sector, composed of a unit measure of identical firms. Firms hire workers to produce final consumption goods.

At  $t = 0$ , each firm is endowed with  $M_0$  units of domestic currency. At time  $t + 1$ , the money supply is given by some  $M_{t+1}$  that is determined by the central bank. At the beginning of each period, the additional units of money are equally distributed among firms by a helicopter. Each firm gets  $T_t$  units of money at time  $t$ , where:

$$T_t = M_t - M_{t-1} \tag{3.22}$$

There is a market for cash. The price of 1 unit of cash at the beginning of period  $t$  is set at  $1 + r_t$  units of output at the end of period  $t$ . Denote by  $Q_{i,t}$  the money holding of firm  $i$  carried over from time  $t - 1$ , and by  $Q_{i,t}^d$  the amount of liquidity demanded by firm  $i$  at the beginning of period  $t$  (given the price  $r_t$ ).

At the beginning of period  $t$  (for  $t \geq 1$ ), firm  $i$ 's money holdings are:

$$M_{i,t} = Q_{i,t} + Q_{i,t}^d + T_t \quad (3.23)$$

There is a cash in advance constraint on hiring labor. In other words, the amount of labor employed by firm  $i$  must satisfy the following condition:

$$w_t L_{i,t}^e \leq M_{i,t} \quad (3.24)$$

Where  $L_{i,t}^e$  denotes the labor employed by firm  $i$  at time  $t$ ,  $M_{i,t}$  is the money holdings of firm  $i$  at time  $t$  and  $w_t$  denotes the nominal wage at time  $t$ .

The nominal price of goods at time  $t$ , denoted  $p_t$ , is taken as given by the firms. Consumers purchase goods with cash at the equilibrium price (implicitly, there is a cash in advance constraint on consumption as well).

As revenues are generated, firms decide whether to hold on to cash in order to be able to make use of it in the next period, or whether to distribute the cash as dividends. In equilibrium, within a period, money will be transferred back and forth between firms and households multiple times: first, firms will transfer their money to households as wage bills. Households will use their wage bills to buy products; this will generate some revenue, which, given the initial high marginal utility of consumption, share holders will decide to distribute as dividends. At the end of the period, when the marginal utility of consumption of shareholders is sufficiently low, firms will decide not to distribute dividends and rather use the revenues as cash reserves.

**Menu costs.** I assume that changing prices is socially costly. In period  $t + 1$ , the output cost of changing the equilibrium price from  $p_t$  to  $p_{t+1}$  is given by  $\kappa(p_t, p_{t+1})$ , where  $\kappa(p_t, p_{t+1}) = \epsilon > 0$  for  $p_t \neq p_{t+1}$  and  $\kappa(p_t, p_{t+1}) = 0$  for  $p_t = p_{t+1}$ . I assume for

simplicity that this cost is incurred by consumers in a lump-sum fashion at time  $t$ .

Assuming some form of menu costs is necessary in order to guarantee that the central bank's optimal policy is to insure price stability. Otherwise, there are multiple equilibria<sup>3</sup>.

**Households.** The utility of households is given by:

$$U(c_0, c_1, c_2 \dots) = \sum_{t=0}^{\infty} \beta^t u(c_t) \quad (3.25)$$

Where  $u(\cdot)$  takes the form:

$$u(c) = \frac{c^{1-\theta}}{1-\theta} \quad (3.26)$$

I assume that  $\beta\rho^{1-\theta} < 1$ . This assumption will be necessary to assure that utility takes a finite value in the autarkic equilibrium (in which consumption equals output and hence grows at a rate  $\rho$ ).

Each household owns one productive firm. I assume that the agent indexed  $i$  owns the firm indexed  $i$ . Households consume their wage bill and their dividends, minus the cost of the price change:

$$p_t c_{i,t} = w_t L + d_{i,t} - \frac{\kappa(p_t, p_{t+1})}{p_t} \quad (3.27)$$

Implicitly, it is assumed that household  $i$  cannot supply its own labor to firm  $i$ . Otherwise, the cash in advance constraint would be nonsensical, as households could supply their own labor to the firms that they own and “pay themselves” later with dividends. The cash in advance constraint captures a reality in which production inputs are differentiated, and production requires the purchase of a variety of production inputs sold in a market setting.

---

<sup>3</sup>In the absence of a menu cost assumption, any sequence of money supply can achieve the first best level. This result changes if we assume that labor supply is elastic. In this case, in the absence of menu costs, the optimal sequence of money supplies implements the Friedman rule (Friedman [1969]).



**The central bank.** The central bank pre-commits to a sequence of money supplies,  $(M_1, M_2, \dots)$  (I assume that  $M_0$  is given). The aim of the central bank is to maximize household utility.

### 3.4 Closed Economy Equilibrium

In the closed economy equilibrium, liquidity constraints are binding in the sense that the marginal product of labor is higher than the real wage. This is because firms will agree to carry cash reserves only if there is a high enough return to money. The liquidity constraints will be more binding for economies that are growing at a faster rate. However, despite the fact that liquidity constraints may be very binding, output and consumption are still at their first best level.

**Definition 3** *An equilibrium of the closed economy is a sequence of prices  $(p_0, p_1, \dots)$ , a sequence of rates of return on liquidity  $(r_0, r_1, \dots)$ , a sequence of nominal wages  $(w_0, w_1, \dots)$ , a sequence of employed labor  $\{(L_{i,0}^e, L_{i,1}^e, \dots)\}_{i \in [0,1]}$ , a sequence of outputs  $(Y_0, Y_1, \dots)$ , a sequence of cash holdings  $\{(Q_{i,0}, Q_{i,1}, \dots)\}_{i \in [0,1]}$ , a sequence of liquidity demands  $\{(Q_{i,0}^d, Q_{i,1}^d, \dots)\}_{i \in [0,1]}$ , a sequence of firms' initial money holdings  $\{(M_{i,0}, M_{i,1}, \dots)\}_{i \in [0,1]}$  a sequence of money supplies  $(M_0, M_1, \dots)$  and a sequence of consumptions  $\{(c_{i,0}, c_{i,1}, \dots)\}_{i \in [0,1]}$  that jointly satisfy the following conditions:*

1. *Given the wage and the price sequences, the consumption sequence, the labor sequence, and the cash sequence solve the optimization problem of agent  $i$ :*

$$\max_{c_{i,t}, L_{i,t}^e, Q_{i,t+1}, Q_{i,t}^d} \sum_{t=0}^{\infty} \beta^t u(c_{i,t}) \quad (3.28)$$

s. t.

$$p_t c_{i,t} = w_t L_{i,t} + d_{i,t} - \frac{\kappa(p_t, p_{t+1})}{p_t} \quad (3.29)$$

$$d_{i,t} = p_t F_t(L_{i,t}^e) + M_{i,t} - w_t L_{i,t}^e - Q_{i,t+1} - (1 + r_t) Q_{i,t}^d \quad (3.30)$$

$$w_t L_{i,t}^e \leq M_{i,t} \quad (3.31)$$

$$M_{i,t+1} = Q_{i,t+1} + Q_{i,t+1}^d + T_{t+1} \quad (3.32)$$

Where  $T_t = M_t - M_{t-1}$ , and  $M_{0,i} = M_0$ .

2. *Goods market clearing condition:*

$$Y_t = \int_0^1 \rho F_t(L_{i,t}^e) di = \int_0^1 c_{i,t} di + \kappa(p_t, p_{t+1}) \quad (3.33)$$

3. *Labor market clearing condition:*

$$\int_0^1 L_{i,t}^e di = L \quad (3.34)$$

4. *Money market clearing condition:*

$$\int_0^1 M_{i,t} di = M_t \quad (3.35)$$

5. *The sequence  $(M_1, M_2, \dots)$  maximizes household utility.*

It will be useful to compare the properties of this equilibrium to the first best of this economy, where the “first best” refers to the unconstrained benchmark without the liquidity constraint and without menu costs.

**Lemma 16** *Consumption is at its first best level if and only if prices are stable.*

**Proof:** If the liquidity constraint is not binding, each firm sets the marginal product of labor to be equal to the wage. If the liquidity constraint is binding, then each firm sets  $\frac{F_t'(L_{i,t}^e)}{w_t} = 1 + r_t$ . Thus, output is at its first best level, as the economy is at full employment and the marginal product of labor is equated across firms. If  $p_t = p_{t-1}$ , then  $\kappa(p_t, p_{t+1}) = 0$  and consumption is equal output; in this case, consumption is at its first best level. If  $p_t \neq p_{t-1}$ , consumers have to incur some

menu costs, so consumption is less than output, and hence less than the first best level.

A corollary of this lemma is that the central bank would like to choose  $(M_1, M_2, \dots)$  in a way that achieves price stability.

The following proposition characterizes the closed economy equilibrium:

**Proposition 9** *There exists a unique equilibrium. In equilibrium, the liquidity constraint is binding for all  $t$ :  $1 + r^{aut} = \frac{p_t F'_t(L)}{w_t} = \frac{\rho^\theta}{\beta}$ .*

The intuition is as follows. The consumer-shareholder's Euler equation is given by:

$$\frac{1}{p_t} u'(c_t) = \frac{\beta}{p_{t+1}} u'(c_{t+1}) \frac{p_{t+1} F'_{t+1}(L)}{w_{t+1}} \quad (3.36)$$

The term  $\frac{p_{t+1} F'_{t+1}(L)}{w_{t+1}}$  is the additional revenue generated in period  $t + 1$  from carrying over one more unit of currency. This unit of currency can hire  $\frac{1}{w_{t+1}}$  units of labor; each unit of labor produces at the margin a revenue of  $p_{t+1} F'_{t+1}(L)$ .

In equilibrium, the central bank chooses a sequence of money supplies that assures price stability. Thus,  $p_t = p_{t+1}$ , and, by Lemma 16,  $c_t = F_t(L)$ . Thus, the Euler equation can be rewritten as:

$$u'(F_t(L)) = \beta u'(F_{t+1}(L)) \frac{p_{t+1} F'_{t+1}(L)}{w_{t+1}} \quad (3.37)$$

From this, it is straightforward to conclude that firms are constrained in equilibrium, in the sense that the marginal product of labor is higher than the wage:

$$\frac{p_{t+1} F'_{t+1}(L)}{w_{t+1}} = \frac{u'(F_t(L))}{\beta u'(F_{t+1}(L))} > 1 \quad (3.38)$$

Thus, firms are constrained in equilibrium and use their entire money holdings to hire labor.

Similar to the model in section 3.2, there is a positive relation between the growth rate of the economy and the equilibrium rate of return on liquidity. From equation

3.38:

$$1 + r^{aut} = \frac{p_{t+1}F'_{t+1}(L)}{w_{t+1}} = \frac{u'(F_t(L))}{\beta u'(\rho F_t(L))} = \frac{F_t(L)^{-\theta}}{\beta(\rho F_t(L))^{-\theta}} = \frac{\rho^\theta}{\beta} \quad (3.39)$$

Faster growing economies (higher  $\rho$ ) are more liquidity constrained in the sense that the ratio of the marginal product of labor and the wage is higher. This is because higher output growth leads to higher consumption growth; the marginal utility of consumption therefore declines more rapidly, and firms require a higher rate of return on money to be willing to carry over money from one period to the next.

### 3.5 Integrated Equilibrium

Consider the following global equilibrium environment. There is a large “developed” economy that is at its steady state level of output ( $\rho = 1$ ). The central bank implements the optimal policy and chooses a constant price level. In this economy, the within period return to holding cash is given by (by equation 3.39 with  $\rho = 1$ ):

$$1 + r = \frac{p_t F'_t(L)}{w_t} = \frac{1}{\beta} \quad (3.40)$$

I normalize the price of goods in terms of international currency to be  $p = 1$ .

In this global environment, I assume a small open economy that grows at rate  $\rho > 1$ . The exchange rate is fixed at 1 and the central bank follows a 100% reserve ratio. Later I will show that abandoning the reserve requirement can increase equilibrium welfare. The initial supply of domestic currency at time 0 is  $M_0$ . At time  $t$ , the amount of domestic currency is determined by the demand for domestic currency given domestic wages and the fixed exchange rate. The price of goods in terms of the foreign currency is constant at  $p_t = 1$ . This implies that the economy incurs no menu costs due to price changes.

I assume that borrowing is limited in the following sense. Agents can borrow and lend only within a period; if agents borrow, they must repay whatever they borrow at the end of the period. This structure will imply that domestic firms will borrow for liquidity purposes, but households will not be able to borrow to smooth consumption

across time.

Given the assumption that the price of goods is constant, the only feasible equilibrium is the constrained equilibrium, in which the between-periods return on holding cash is 0, and the within-period return on holding cash is  $1 + r$ :

$$\frac{F'_t(L)}{w_t} = 1 + r \quad (3.41)$$

**Definition 4** *An equilibrium of the integrated economy is a sequence of nominal wages  $(w_0, w_1, \dots)$ , a sequence of employed labor  $\{(L_{i,0}^e, L_{i,1}^e, \dots)\}_{i \in [0,1]}$ , a sequence of outputs  $(Y_0, Y_1, \dots)$ , a sequence of domestic cash holdings  $\{(Q_{i,0}, Q_{i,1}, \dots)\}_{i \in [0,1]}$ , a sequence of liquidity demands  $\{(Q_{i,0}^d, Q_{i,1}^d, \dots)\}_{i \in [0,1]}$ , a sequence of foreign liquidity supply  $(Q_0^*, Q_1^*, \dots)$  and a sequence of consumptions  $\{(c_{i,0}, c_{i,1}, \dots)\}_{i \in [0,1]}$  that jointly satisfy the following conditions:*

1. *Given the wage sequence, the consumption sequence, the labor sequence, and the cash sequence solve:*

$$\max_{c_{i,t}, L_{i,t}^e, Q_{i,t+1}, Q_{i,t}^d} \sum_{t=0}^{\infty} \beta^t u(c_{i,t}) \quad (3.42)$$

s.t.

$$c_{i,t} = w_t L + d_{i,t} \quad (3.43)$$

$$d_{i,t} = F_t(L_{i,t}^e) + Q_{i,t} + Q_{i,t}^d - w_t L_{i,t}^e - Q_{i,t+1} - (1 + r)Q_{i,t}^d \quad (3.44)$$

$$w_t L_{i,t}^e \leq Q_{i,t} + Q_{i,t}^d \quad (3.45)$$

$Q_{i,0} = M_0$  is given.

2. *Goods market clearing condition:*

$$Y_t = \int_0^1 \rho F_t(L_{i,t}^e) di = \int_0^1 c_{i,t} di + rQ_t^* \quad (3.46)$$

3. *Labor market clearing condition:*

$$\int_0^1 L_{i,t}^e di = L \quad (3.47)$$

4. *Liquidity market clearing condition:*

$$\int_0^1 Q_{i,t}^d di = Q_t^* \quad (3.48)$$

**Proposition 10** *For  $M_0$  sufficiently small:*

1. *In the financially integrated economy, firms always hold 0 cash reserves.*
2. *Financial integration reduces equilibrium consumption:  $c_t < c_t^{aut}$  for all  $t$ .*

Note that firms will agree to hold cash only as long as:

$$u'(c_t) \leq \beta u'(c_{t+1})(1+r) = u'(c_{t+1}) \quad (3.49)$$

It turns out that this condition is violated in equilibrium, as the economy experiences positive consumption growth which it is unable to smooth. Firms therefore choose to hold 0 cash reserves, and are always constrained in equilibrium. As in the simple example in section 3.2, financial integration creates a dependence on foreign liquidity supply as the wage bill appreciates. In this dynamic setting, there is an additional effect: not only are wages higher, but the equilibrium choice of cash reserves is lower, further increasing the reliance of foreign liquidity supply and the equilibrium rents absorbed by foreign liquidity suppliers.

As in the simple example, the equilibrium inefficiency results from the fact that constrained entrepreneurs do not internalize the effects of their borrowing decisions on the equilibrium wage. From the perspective of each firm, the wage is depressed so borrowing from foreigners is profitable. However, collectively, firms would be better off if borrowing from foreigners was restricted. Dividends would be higher, both because wages would be lower and because there would be no payment to foreign liquidity suppliers.

From the worker's perspective, financial integration is beneficial, as it increases the real wage. In this model, the negative effect on dividends more than offsets this effect, resulting in overall lower consumption. This model therefore suggests that financial integration is associated with a redistribution of surplus from capital owners (firms) to workers, that cannot be corrected by redistribution.

### 3.6 Global Imbalances

In the integrated equilibrium, despite the fact that agents would like to borrow, the central bank accumulates foreign reserves over time, and the economy runs a trade deficit.

The current account is given by:

$$CA_t = (1 + r)Q_t^* - Q_t^* = rQ_t^* \quad (3.50)$$

Note that:

$$\frac{F'_t(L)}{w_t} = 1 + r \Rightarrow \frac{F'_t(L)}{1 + r} = w_t \Rightarrow \frac{F'_t(L)L}{1 + r} = w_t L \quad (3.51)$$

As foreign liquidity supply finances the entire wage bill. by equation 3.51,

$$w_t L = Q_t^* \Rightarrow \frac{F'_t(L)L}{1 + r} = w_t L = Q_t^* \quad (3.52)$$

The current account in period  $t$  is therefore:

$$CA_t = \frac{rF'_t(L)L}{1 + r} \quad (3.53)$$

Assuming the functional form  $F_t = AL^\alpha$ , we have that  $F'_t(L)L = \alpha F_t(L)$ . Rewriting the equation above,

$$CA_t = \frac{r\alpha F_t(L)}{1 + r} \Rightarrow \frac{CA_t}{F_t(L)} = \frac{r\alpha}{1 + r} \quad (3.54)$$

To calibrate the current account surplus implied by this equation, I choose  $r$  to be equal to the US prime rate (taken from the St. Louis Fed's FRED database) plus 3%. This choice is roughly consistent with the return on risky productive lending. The choice of  $\alpha$  is more tricky:  $\alpha$  need not be interpreted necessarily as the labor share, but rather as the share of inputs that must be purchased in advance of production. I choose  $\alpha = 0.6$ , based on estimates of the credit share in production in Evans et al. [2002]<sup>4</sup>.

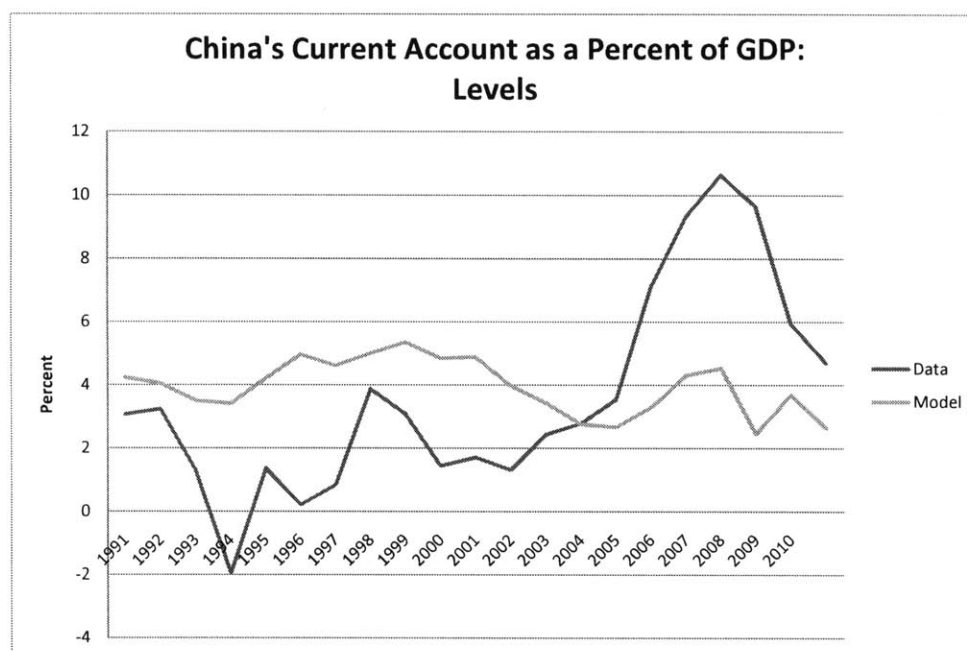


Figure 3-1: The Chinese current account as a percent of gdp.

Figure 3-1 plots the current account predicted by the model and the actual Chinese current account between the years 1990 and 2010 (both as a percent of GDP). The model predicts a current account surplus of between 2%-6%. This range falls well

<sup>4</sup>Evans et al. [2002] use a panel of 82 countries covering 21 years to estimate a translog production function, and find the share of credit to be around 0.6. An alternative estimate can be found in Khan and Ahmad [1985], that estimates the share of money in the production function to be 0.43 in the manufacturing sector in Pakistan.



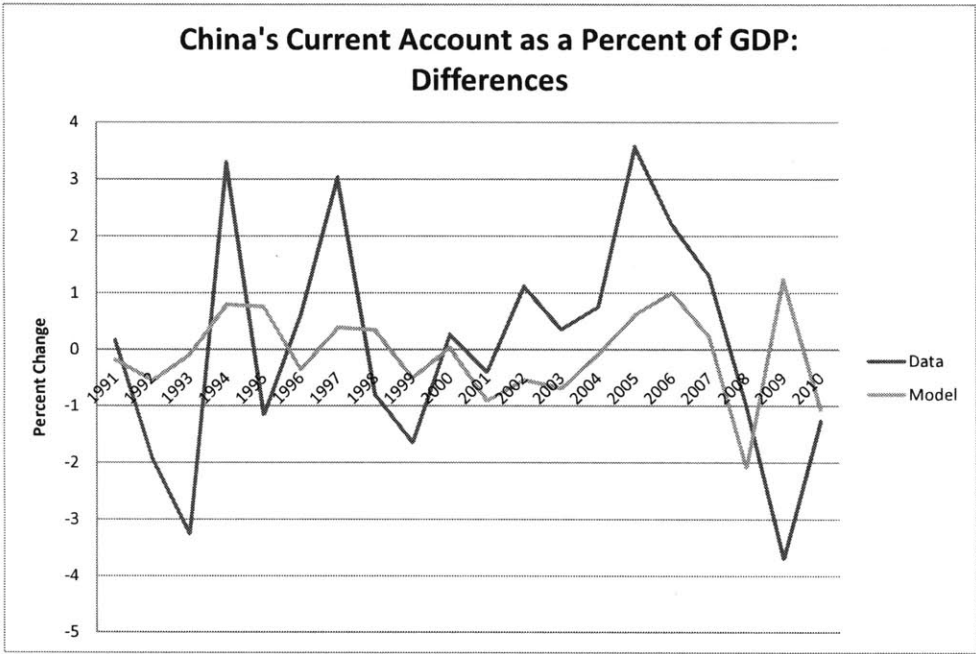


Figure 3-2: The Chinese current account as a percent of gdp: first differences on an annual frequency.

within the range of the actual Chinese current account, which is between -2% and 10%. It is possible that the larger range can be accounted for by time varying risk premia on Chinese short-term borrowing<sup>5</sup>. Note further that, on a yearly frequency, the increases in the current account predicted by the model roughly coincide with those observed in the data (see figure 3-2). Of course, on a lower frequency, the relationship seems to be failing as the Chinese current account is increasing while the current account predicted by the model is slightly downward trending.

### 3.6.1 Optimal Policy

Interestingly, the central bank accumulates foreign reserves, despite the fact that domestic consumers would like to borrow. To see this, note that the demand for domestic currency is increasing over time: the demand for domestic currency at time  $t$  is given by  $Q_t^*$ , which, from the analysis above, is increasing at a rate  $\rho$ . The central bank therefore has to print new currency every period, and accumulates foreign reserves.

Note that the analysis above was done under the assumption that the central bank must follow a fixed exchange rate regime with a reserve ratio of 1. A central bank which lacks credibility and is vulnerable to attacks will therefore be forced to accumulate cash reserves. The central bank may choose to use these cash reserves to buy liquid low-risk assets such as treasuries; the interest on treasuries can be distributed to domestic agents immediately without compromising the reserve ratio. Will the treasuries ever be redeemed? This depends on whether the central bank continues to be vulnerable to attacks forever. It is possible that over time, the central bank will acquire credibility and will no longer need to back its currency with the international currency. At that point, the central bank will find it optimal to stop accumulating bonds. In this sense, the US current account deficit can be seen as “unsustainable”.

---

<sup>5</sup>See Broner et al. [Forthcoming] for evidence that time varying risk premia plays an important role in determining emerging market borrowing rates. This mechanism may potentially help account for the large discrepancy between the model and the data during the recent crisis, in which interest rates were held very low, but firms experienced difficulty acquiring credit. The short term borrowing rate faced by Chinese firms may be badly proxied by the US prime lending rate.

If, at some point, the Chinese government acquires sufficient credibility to have an unbacked currency, the demand for US treasuries will decline.

Two questions arise: first, can the central bank increase welfare by abandoning the commitment to a 100% reserve ratio? Second, can the central bank further improve welfare by moving to a flexible exchange rate regime?

The answer to the first question is yes: the central bank can increase welfare if it is able to commit to a fixed exchange rate without a 100% reserve ratio. To see this, note that given a fixed exchange rate, the amount of domestic currency at time  $t$  must be equal to the demand for domestic currency in the 100% reserve ratio case analyzed above, denoted  $\hat{Q}_t^*$ :

$$\hat{Q}_t^* = \frac{F'_t(L)L}{1+r} \quad (3.55)$$

The amount of domestic currency therefore must grow at a rate  $\rho$ . Assume that the only restriction on the central bank is that it must have a positive reserve ratio (it cannot hold a negative amount of foreign currency). At the end of period  $t$ , foreigners hold  $\hat{Q}_t^*$  units of domestic currency. At the beginning of time  $t+1$ , the central bank can improve welfare by distributing the new units of currency directly to domestic firms, without backing it by foreign reserves.

Firms will receive  $\hat{Q}_{t+1}^* - \hat{Q}_t^*$  units of currency, thereby absorbing a fraction  $\frac{\rho-1}{\rho}$  of the rents to liquidity supply at time  $t+1$ . By following this policy every period, the government can transfer  $\frac{\rho-1}{\rho}$  of the rents to liquidity supply to domestic agents. Note that this policy is consistent with a fixed exchange rate, as all equilibrium prices and rates of return replicate the 100% reserve ratio case analyzed in the previous section. Foreigners are therefore still willing to trade one unit of domestic currency for one unit of foreign currency.

Note that under this policy, the reserve ratio will approach 0 as  $t \rightarrow \infty$ . To see this, note that if the central bank begins to implement this policy at time 1, the

foreign reserves at any time  $t$  remain at  $M_0$ , so the reserve ratio is:

$$RR = \frac{M_0}{\hat{Q}_t^*} = \frac{M_0}{\frac{F_t'(L)L}{1+r}} = \frac{M_0}{\frac{\rho^t F_0'(L)L}{1+r}} \rightarrow_{t \rightarrow \infty} 0 \quad (3.56)$$

Thus, a central bank trying to improve welfare by following this policy must be sufficiently credible to be able to sustain a fixed exchange rate with effectively a 0 reserve ratio.

Can the central bank improve welfare further by abandoning the fixed exchange rate regime and moving to a flexible exchange rate regime? Unfortunately, no. The reason is as follows. For foreigners to be willing to hold a finite amount of domestic currency between the end of period  $t$  and the beginning of period  $t + 1$ , it must be the case that the exchange rate remains the same between these two dates. If it is expected that the central bank will depreciate the currency, nobody will hold domestic currency, as it is expected that it will be cheaper to purchase domestic currency at the beginning of the next period. Similarly, if it is expected that the central bank will appreciate the currency, everyone will want to hold domestic currency. The exchange rate at the beginning of period  $t$  therefore must be equal to the exchange rate at the end of the previous period.

Similarly, it must be the case that the within-period return on each unit of domestic currency is  $1 + r$ , by the international no arbitrage condition. Thus, if foreigners hold  $Q_t^*$  units of domestic currency at the beginning of time  $t + 1$ , by the end of period  $t + 1$  the economy must pay at least  $(1 + r)Q_t^*$  units of output as rents to foreigners. The best the central bank can do is therefore replicate the allocation of the fixed exchange rate regime with no reserve requirement, and distribute the new units of domestic currency directly to firms.

### 3.7 Instability

As the wage bill increases following financial integration, the economy becomes increasingly reliant on foreign liquidity supply. Domestic firms are no longer able to

afford the wage bill with their own money holdings. In the presence of some wage stickiness, this means that the economy is particularly vulnerable to shocks to foreign liquidity supply<sup>6</sup>.

I augment the model to allow for wage stickiness. Assume that at the time of financial integration, employers and workers contract on a sequence of wages,  $\hat{w}_t, \hat{w}_{t+1}, \dots$ . The contracted wages are determined by the “no shock” equilibrium described in section 3.5. Thus, the wage  $\hat{w}_t$  is characterized by the following international no-arbitrage condition, that takes into account full employment of labor:

$$\frac{F'_t(L)}{\hat{w}_t} = 1 + r \quad (3.57)$$

Assume that there is a single period shock to the foreigners’ willingness to supply liquidity that leads the within-period borrowing rate to jump from  $r$  to  $r' > r$ .

As wages are sticky, this increase in the interest rate necessarily implies a drop in production. The international no arbitrage condition is now given by:

$$\frac{F'_t(L_t^e)}{\hat{w}_t} = 1 + r' > 1 + r = \frac{F'_t(L)}{\hat{w}_t} \Rightarrow L_t^e < L \quad (3.58)$$

Thus, unemployment increases and output drops. Consumption at time  $t$  drops as well:

**Lemma 17** *Consumption drops following an increase in  $r$ .*

It is easy to see that  $Q_{i,t+1} = 0$ , as borrowing conditions at  $t + 1$  are expected to go back to normal; firms therefore have no incentive to save, and will choose to hold 0 cash reserves. Thus, consumption returns back to the no-shock equilibrium benchmark. As consumption decreases in period  $t$  and returns to normal thereafter, it can be concluded that the shock to the foreigners’ willingness to lend reduces equilibrium welfare.

---

<sup>6</sup>This is potentially a realistic concern, as the literature suggests that foreign liquidity supply to emerging economies is highly volatile (see Fostel and Geanakoplos [2008] and Eden [2011a] for models).

**“Beggar thy neighbor” policies.** Given a shock to the foreigners willingness to lend, the government may have an incentive to depreciate the currency in order to boost employment. To see this, note that the domestic return to a unit of foreign currency is determined by the exchange rate  $e_t$ . A unit of labor costs  $\frac{w_t}{e_t}$  units of foreign currency. The return to foreign currency is therefore equal to:

$$\frac{F'_t(L_t^e)}{\frac{w_t}{e_t}} = e_t \frac{F'_t(L_t^e)}{w_t} \quad (3.59)$$

Consider a government that sets the exchange rate at  $e_t$ . Employment is then given by the following international no arbitrage condition:

$$1 + r' = e_t \frac{F'_t(L_t^e)}{\hat{w}_t} \quad (3.60)$$

**Lemma 18** *Consumption at time  $t$  is maximized when  $e_t$  is chosen such that full employment is restored.*

Given a crisis, the central bank faces a tradeoff: on the one hand, a depreciation may compromise the central bank’s credibility. This is problematic as there are welfare gains from the ability to institute a currency that is not backed by foreign reserves. On the other hand, during a crisis, there is an incentive to renege on the promise to hold the exchange rate fixed, as a depreciation can help restore full employment.

### 3.8 Conclusion

The presence of binding liquidity constraints implies a transfer of surplus to liquidity suppliers. In the closed economy, liquidity is supplied domestically so the presence of liquidity constraints is welfare-neutral. However, in the integrated equilibrium, binding liquidity constraints in a fast-growing economy imply a transfer of surplus to foreign liquidity suppliers, thereby reducing domestic equilibrium welfare.

From a policy perspective, this model suggests that emerging economies have an

incentive to discourage foreign liquidity supply, even if domestic entrepreneurs are heavily constrained. Opening to capital flows will increase the supply of liquidity, but this will only bid up the domestic input prices such as labor and land. This, in turn, will make the domestic firms less liquidity constrained but more heavily reliant on foreign liquidity supply. In a dynamic equilibrium, the lower return on holding liquidity will imply that firms choose not to hold sufficient cash reserves to finance their expenses, and rely heavily on foreigners to supply liquidity.

The integrated economy is therefore highly vulnerable to shocks to foreign liquidity supply. In the presence of sticky input prices, a sudden shock to foreigner's willingness to supply liquidity will result in an increase in unemployment and a drop in output. The government therefore has an incentive to depreciate the domestic currency, thereby providing a higher dollar rate of return and encouraging foreign liquidity supply.

This temptation to depreciate the currency creates a certain tension. A central bank that is sufficiently credible to maintain a fixed exchange rate that is not vulnerable to attacks can increase domestic welfare, as it can distribute new currency directly to domestic firms without backing it by foreign reserves. As long as this is done in a manner consistent with the fixed exchange rate, it allows domestic firms to absorb some of the rents to liquidity supply. A shock to foreigner's willingness to supply liquidity therefore creates a tension between the short term gains of a depreciation and the long term gains of central bank credibility.

An important thing to note is that, in this model, the driver of global imbalances is the differences in growth rates: the fast-growing economy is liquidity constrained in equilibrium, and the steady-state developed world is the equilibrium liquidity supplier. This suggests that global imbalances are a temporary phenomenon: once the emerging markets converge to the steady state growth level, we should see a balanced current account. This suggests that, from the developed market's perspective, the rents from supplying liquidity to emerging economies should be viewed as a "temporary" source of income.





# Bibliography

- Andrew B. Abel. Dynamic accumulation of capital accumulation in a cash-in-advance model. *Journal of Monetary Economics*, 16:55–71, 1985.
- S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, August 1994.
- Pol Antras and Ricardo J. Caballero. Trade and capital flows: A financial frictions perspective. *Journal of Political Economy*, 117(4):701–744, August 2009.
- Jagdish N. Bhagwati and Richard A. Brecher. National welfare in an open economy in the presence of foreign-owned factors of production. *Journal of International Economics*, 10:103–115, 1980.
- Fernando A. Broner, Guido Lorenzoni, and Sergio L. Schmukler. Why do emerging economies borrow short term? *Journal of the European Economic Association*, Forthcoming. Forthcoming.
- Ricardo J. Caballero, Emmanuel Farhi, and Pierre-Olivier Gourinchas. An equilibrium model of ‘global imbalances’ and low interest rates. *American Economic Review*, 98(1):358–393, 2008.
- Maya Eden. Financial distortions and the distribution of global volatility. 2011a.
- Maya Eden. The inefficiency of financial intermediation in general equilibrium. 2011b.
- Alun Dwyfur Evans, Christopher J. Green, and Victor Murinde. Human capital and financial development in economic growth: New evidence using the translog

production function. *International Journal of Finance and Economics*, 7:123–140, 2002.

Ana Fostel and John Geanakoplos. Leverage cycles and the anxious economy. *American Economic Review*, 98(4):1211–1244, September 2008.

Milton Friedman. The optimum quantity of money. In *The Optimum Quantity of Money and Other Essays*, pages 1–50. Chicago: Aldine, 1969.

Bengt Holmstrom and Jean Tirole. Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics*, 112(3):663–691, August 1997.

Ashfaq H. Khan and Mushtaq Ahmad. Real money balances in the production function of a developing country. *The Review of Economics and Statistics*, 67(2): 336–340, May 1985.

Anton Korinek. Hot money and serial financial crises. 2010.

Pablo A. Neumeyer and Fabrizio Perri. Business cycles in emerging economies: the role of interest rates. *Journal of Monetary Economics*, 52(2):345–380, 2005.

Alan C. Stockman. Anticipated inflation and the capital stock in a cash-in-advance economy. *Journal of Monetary Economics*, 8:387–393, 1981.

## 3.9 Appendix

### 3.9.1 Proof of Lemma 15

If the liquidity constraint is not binding, firms set the marginal product of labor equal to the wage. As in the unconstrained benchmark, the only rate of return on liquidity that is consistent with market clearing is  $r^{aut} = 0$ .

Assume next that the liquidity constraint is binding, that is, the weak inequality in equation 3.7 holds with equality:

$$wL_i^e = Q + Q_i^d \tag{3.61}$$

The firm's optimization problem (equation 3.6) can be rewritten as:

$$\max_{Q_i^d} \rho F\left(\frac{Q + Q_i^d}{w}\right) - (1 + r)Q_i^d \quad (3.62)$$

The first order condition with respect to  $Q_i^d$  is:

$$\frac{F'(L_i^e)}{w} = 1 + r \quad (3.63)$$

It follows that  $r > 0$ : if  $r = 0$ , the marginal product of labor is equated with the wage and the liquidity constraint is not binding. If  $r < 0$ , the marginal product of labor exceeds the wage and this is never optimal.

It is left to show that when  $r^{aut} > 0$ , the liquidity constraint is binding. This follows from the fact that if the liquidity constraint is not binding,  $r^{aut} = 0$  is the only return consistent with market clearing in the liquidity market.

### 3.9.2 Proof of Proposition 9

In equilibrium, firms must be indifferent, given the price level, between holding on to cash and distributing cash as dividends. This can imply one of two equations, depending on whether the liquidity constraint is binding at  $t + 1$ . I will consider first the situation in which the liquidity constraint is binding. Substituting in  $w_{t+1}L_{t+1} = Q_{t+1} + T_{t+1}$ , the Euler equation implies:

$$\frac{1}{p_t} u'(c_t) = \beta \frac{1}{p_{t+1}} u'(c_{t+1}) \frac{p_{t+1} F'_{t+1}(L_t)}{w_{t+1}} \quad (3.64)$$

The left hand side is the marginal utility of households (share holders) generated from receiving an additional unit of currency in this period. An additional unit of currency can purchase  $\frac{1}{p_t}$  consumption goods, and the marginal utility of consumption is  $u'(c_t)$ . The right hand side represents the marginal utility of households generated from having firms hold an additional unit of cash as reserves. With an additional unit of cash, firms can hire  $\frac{1}{w_{t+1}}$  more workers. Each worker generates an additional

nominal revenue of  $p_{t+1}F'_{t+1}(L)$ . Each unit of cash in period  $t+1$  can finance  $\frac{1}{p_{t+1}}$  units of consumption, and the marginal utility of consumption is  $u'(c_{t+1})$ . The marginal utility of consumption next period is discounted by  $\beta$ .

I conjecture that there exists a sequence  $(M_1, M_2, \dots)$  that guarantees price stability. Under this conjecture, by Lemma 16, output is equal consumption and is at its first best level. Substituting in the market clearing condition,  $c_t = Y_t$ , and the condition that output is at its first best level,

$$\frac{1}{p_t}u'(Y_t) = \beta u'(Y_{t+1}) \frac{F'_{t+1}(L)}{w_{t+1}} = \beta u'(\rho Y_t) \frac{\rho F'_t(L)}{w_{t+1}} \quad (3.65)$$

Substituting in the functional form for  $u(\cdot)$ , after some simple algebraic manipulations:

$$\frac{p_t F'_{t+1}(L)}{w_{t+1}} = \frac{\rho^\theta}{\beta} \quad (3.66)$$

Note that in this environment,  $w_{t+1} = \frac{M_{t+1}}{L}$ . Thus,

$$p_t = \frac{\rho^\theta M_{t+1}}{\beta F'_{t+1}(L)L} = \frac{\rho^\theta M_{t+1}}{\beta \rho F'_t(L)L} \quad (3.67)$$

Thus, for price stability,  $p_t = p_{t+1}$ , and  $M_{t+1} = \rho M_t$ . Thus, the unique price sequence  $M_t = \rho^t M_0$  guarantees price stability, and, from Lemma 16, this will be the central bank's choice.

I have proved therefore that there is a unique equilibrium in which the liquidity constraint is binding. It is left to show that there are no additional equilibria in which the liquidity constraint does not bind. If the liquidity constraint is not binding at  $t+1$ , the Euler equation implies:

$$\frac{1}{p_t}u'(c_t) = \frac{1}{p_{t+1}}\beta u'(c_{t+1}) \quad (3.68)$$

First, I show that given the price sequence  $M_t = \rho^t M_0$ , the unique equilibrium is the constrained equilibrium. Then, I show that any unconstrained market equilibrium is one in which welfare is suboptimal. I conclude that only way that the central bank can implement the first best outcome is by choosing the money supply sequence

$M_t = \rho^t M_0$ . In this case, there is a unique equilibrium in which consumption is at its first best.

Assume therefore that  $M_t = \rho^t$ . Assume in the way of contradiction that there exists an unconstrained equilibrium in which there is price stability. Substituting in the market clearing condition  $c_t = Y_t$ , and using the expressions for utility and output growth, the following relation emerges:

$$\frac{p_{t+1}}{p_t} = \frac{\beta}{\rho^\theta} \quad (3.69)$$

Thus, price stability must be violated. It is left to show that there is no market equilibrium in which there is no price stability. Assuming no price stability, the market clearing condition implies that:

$$c_t = Y_t - \kappa(p_t, p_{t-1}) = Y_t - \epsilon \quad (3.70)$$

Thus, from the Euler equation,

$$\frac{1}{p_t} u'(\rho^t F_0(L) - \epsilon) = \frac{\beta}{p_{t+1}} u'(\rho^{t+1} F_0(L) - \epsilon) \quad (3.71)$$

$$\frac{p_{t+1}}{p_t} = \beta \frac{u'(\rho^{t+1} F_0(L) - \epsilon)}{u'(\rho^t F_0(L) - \epsilon)} \quad (3.72)$$

Denote:

$$\theta = \beta \frac{u'(\rho^{t+1} F_0(L) - \epsilon)}{u'(\rho^t F_0(L) - \epsilon)} \quad (3.73)$$

Note that  $\theta < 1$  and  $p_{t+1} = \theta p_t$ . Thus, real money balances ( $\frac{M_t}{p_t}$ ) grow at rate  $\frac{\rho}{\theta}$ , whereas equilibrium consumption, for  $t$  large, grows at a rate equal approximately to  $\rho$ . This is in violation of the transversality condition: agents accumulate money reserves that they never use. From an individual standpoint, they can do better by consuming more in every period and holding less money reserves. Thus, an unconstrained equilibrium in which money grows at rate  $\rho$  does not exist.

It is left to show that no unconstrained equilibrium can achieve the first best

level of consumption. To see this, note that by equation 3.69, any unconstrained equilibrium in which output is at its first best would require price deflation. By Lemma 16, an economy with price deflation cannot be at its first best.

### 3.9.3 Proof of Proposition 10

Whether or not firms are borrowers or lenders at  $t = 0$  depends on their liquidity endowment. Specifically, firms are lenders if the following condition holds:

$$\frac{F'_0(L)}{(w_0 = \frac{M_0}{L})} > 1 + r \quad (3.74)$$

Rewriting this condition,

$$\frac{F'_0(L)L}{M_0} > 1 + r \quad (3.75)$$

This condition will be true for  $M_0$  sufficiently low. Note that here the liquidity endowment  $M_0$  has a real interpretation: the liquidity endowment is the amount of foreign goods that domestic agents can purchase at time 0. For the firms to be net borrowers, this must be small in comparison to the value of international goods that they can produce. I am going to assume throughout that this is the case. Note that this condition holds in the autarkic constrained equilibrium<sup>7</sup>.

Denote by  $Q_t^*$  the amount of foreign borrowing of firms at time  $t$ . Consumption at time  $t$  is given by:

$$c_t = w_t L + d_t = w_t L + F_t(L) + Q_t - w_t L - rQ_t^* - Q_{t+1} \quad (3.76)$$

$$c_t = F_t(L) + Q_t - rQ_t^* - Q_{t+1} \quad (3.77)$$

What is  $Q_t^*$ ? If the liquidity constraint binding,  $Q_t^* + Q_t = w_t L$ . Thus,  $Q_t^* = w_t L - Q_t$ . Consumption can be rewritten as:

$$c_t = F_t(L) + (1 + r)Q_t - r w_t L - Q_{t+1} \quad (3.78)$$

---

<sup>7</sup>In the autarkic constrained equilibrium, by equation 3.39,  $\frac{p_t F'_t(L)}{w_t} = \frac{\rho^\theta}{\beta}$ . As  $1 + r = \frac{1}{\beta}$ ,  $\frac{p_t F'_t(L)}{w_t} = \rho^\theta(1 + r) > 1 + r$

Note that:

$$\frac{F'_t(L)}{w_t} = (1+r) \Rightarrow \frac{F'_t(L)}{1+r} = w_t \Rightarrow \frac{F'_t(L)L}{1+r} = w_t L \quad (3.79)$$

Thus,

$$c_t = F_t(L) + (1+r)Q_t - \frac{rF'_t(L)L}{1+r} - Q_{t+1} \quad (3.80)$$

For  $t = 0$ , consumption is maximized for  $Q_1 = 0$ . I conjecture that  $Q_t = 0$  for every  $t$  is an optimal choice whenever  $Q_0 = M_0$  is sufficiently low. Under this conjecture,

$$c_1 = F_1(L) - \frac{rF'_1(L)L}{1+r} = \rho(F_0(L) - \frac{rF'_0(L)L}{1+r}) \quad (3.81)$$

For  $M_0$  sufficiently small,

$$c_0 = F_0(L) + (1+r)M_0 - \frac{rF'_0(L)L}{1+r} \leq \rho(F_0(L) - \frac{rF'_0(L)L}{1+r}) = c_1 \quad (3.82)$$

Thus,  $u'(c_0) > u'(c_1)$ , so firms are choosing optimally not to hold any cash and to distribute all sales revenues as dividends. For periods  $t \geq 1$ , not holding cash continues to be optimal, as consumption is given by:

$$c_t = F_t(L) - \frac{rF'_t(L)L}{1+r} < \rho(F_t(L) - \frac{rF'_t(L)L}{1+r}) = c_{t+1} \quad (3.83)$$

Thus, in equilibrium, firms hold 0 cash reserves in all period, so the second part of the proposition is proven.

The first part of the proposition, that consumption is lower for all  $t$ , is trivial for  $t \geq 1$ , as:

$$c_t = F_t(L) - \frac{rF'_t(L)L}{1+r} < F_t(L) = c_t^{aut} \quad (3.84)$$

For  $t = 0$ , this is true for a sufficiently small  $Q_0$ :

$$c_0 = F_0(L) + (1+r)Q_0 - \frac{rF'_0(L)L}{1+r} \leq F_0(L) = c_0^{aut} \quad (3.85)$$

### 3.9.4 Proof of Lemma 17

Assume that  $r' > r$ , and that the wage is fixed at its “no shock” level  $\hat{w}_t$  such that  $\frac{F'_t(L)}{\hat{w}_t} = 1 + r$ . Consumption after the shock is given by:

$$c_t = w_t L_t^e + F_t(L_t^e) - w_t L^e - Q_{t+1} - r' Q_t^* \leq F_t(L_t^e) - r' Q_t^* \quad (3.86)$$

$$= F_t(L_t^e) - r' \hat{w}_t L_t^e < F_t(L_t^e) - r \hat{w}_t L_t^e \quad (3.87)$$

Consider the function:

$$\phi(l) = F_t(l) - r \hat{w}_t l \quad (3.88)$$

Thus, we have that:

$$c_t < \phi(L_t^e) \quad (3.89)$$

The maximum value that the function  $\phi(\cdot)$  takes is at  $l^{\max}$ , that is determined by the first order condition:

$$\phi'(l) = 0 \Rightarrow F'_t(l^{\max}) = r \hat{w}_t \quad (3.90)$$

Note further that  $\phi''(l^{\max})$  is negative:

$$\phi''(l^{\max}) = F''_t(l^{\max}) - r \hat{w}_t < 0 \quad (3.91)$$

Thus, for  $l < l^{\max}$ ,  $\phi(\cdot)$  is increasing in  $l$ . Note that  $L < l^{\max}$ , as:

$$F'_t(L) = (1 + r) \hat{w}_t > r \hat{w}_t = F'_t(l^{\max}) \Rightarrow L < l^{\max} \quad (3.92)$$

As  $L > L_t^e$ , both  $L$  and  $L_t^e$  are less than  $l^{\max}$ . As  $\phi(\cdot)$  is increasing in this range,

$$c_t < \phi(L_t^e) < \phi(L) = F_t(L) - r \hat{w}_t L = \hat{c}_t \quad (3.93)$$

Where  $\hat{c}_t$  is the level of consumption that domestic consumers would have in the absence of the shock (that is, if the  $r$  was at its expected level).



### 3.9.5 Proof of Lemma 18

Recall that the international no arbitrage condition is given by:

$$1 + r' = e_t \frac{F'_t(L_t^e)}{\hat{w}_t} \quad (3.94)$$

Note that a higher value of  $e_t$  corresponds to higher employment.

Let  $\tilde{r}$  be given by:

$$1 + \tilde{r} = \frac{1 + r'}{e_t} \quad (3.95)$$

By choosing  $e_t$ , the government can control  $\tilde{r}$ . Note that this is the real rate of return that firms have to pay liquidity suppliers in equilibrium:

$$1 + \tilde{r} = \frac{F'_t(L_t^e)}{\hat{w}_t} \quad (3.96)$$

Consumption is given by:

$$c_t = F_t(L_t^e) - \tilde{r}Q_t^* = F_t(L_t^e) - \tilde{r}\hat{w}_tL_t^e \quad (3.97)$$

$$= F_t(L_t^e) - \left(\frac{F'_t(L_t^e)}{\hat{w}_t} - 1\right)\hat{w}_tL_t^e = F_t(L_t^e) - F'_t(L_t^e)L_t^e + \hat{w}_tL_t^e \quad (3.98)$$

Define the function  $\psi(\cdot)$  as:

$$\psi(l) = F_t(l) - F'_t(l)l + \hat{w}_tl \quad (3.99)$$

The derivative of  $\psi(\cdot)$  with respect to  $l$  is always positive:

$$\psi'(l) = F'_t(l) - F''_t(l)l - F'_t(l) + \hat{w}_t = -F''_t(l)l + \hat{w}_t > 0 \quad (3.100)$$

Thus, it is optimal for the government to set  $e_t$  such that the economy is at full employment.

