

9-2012

The Sources of Financial Profit: A Theoretical and Empirical Investigation of the Transformation of Banking in the US

Irene G. Levina

University of Massachusetts Amherst, ilevina07@gmail.com

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations



Part of the [Economics Commons](#)

Recommended Citation

Levina, Irene G., "The Sources of Financial Profit: A Theoretical and Empirical Investigation of the Transformation of Banking in the US" (2012). *Open Access Dissertations*. 616.

https://scholarworks.umass.edu/open_access_dissertations/616

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

THE SOURCES OF FINANCIAL PROFIT: A THEORETICAL AND EMPIRICAL
INVESTIGATION OF THE TRANSFORMATION OF BANKING IN THE US

A Dissertation Presented

by

IREN G. LEVINA

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2012

Economics

© Copyright by Iren G. Levina 2012

All Rights Reserved

THE SOURCES OF FINANCIAL PROFIT: A THEORETICAL AND EMPIRICAL
INVESTIGATION OF THE TRANSFORMATION OF BANKING IN THE US

A Dissertation Presented

by

IREN G. LEVINA

Approved as to style and content by:

David Kotz, Chair

Gerald Epstein, Member

Dan Clawson, Outside Member

Michael Ash, Department Chair
Department of Economics

ACKNOWLEDGMENTS

I would like to thank my supervisor, David Kotz, for invaluable help, thought-provoking comments and encouragement on all stages of dissertation writing. Feedback from Jerry Epstein and Dan Clawson has helped improve and clarify the arguments of the present study. I am also grateful to Costas Lapavitsas who has been extremely generous with his knowledge. Discussions with Costas have inspired, challenged my thinking, and helped develop a theoretical framework of this study.

I am thankful to Bob Pollin who taught a class on Money, Credit and Financial Markets during my second semester at UMass. A research question that lies at the heart of this dissertation, namely, the sources of financial profit, came out of discussions in this class. Duncan Foley offered helpful feedback on various parts of the work. This dissertation has also largely benefited from discussions with Hasan Comert, Duncan Lindo, Josh Mason, Tomas Rotta, and Paulo dos Santos.

This work would not have been possible without my parents who fostered my intellectual curiosity and capacity for hard work since early childhood. Dale Melcher and Bill Newman have become my American parents and made me feel at home for the years of my graduate studies. Special thanks to Alyssa Schneebaum, Xiao Jiang, Riko Rosete, and Jenna Allard who shared the moments of joy and inspiration with me, encouraged and stood by me in challenging times, and showed me the meaning of true friendship.

ABSTRACT

THE SOURCES OF FINANCIAL PROFIT: A THEORETICAL AND EMPIRICAL
INVESTIGATION OF THE TRANSFORMATION OF BANKING IN THE US

SEPTEMBER 2012

IREN G. LEVINA, B.A., MOSCOW STATE UNIVERSITY

M.A., MOSCOW STATE UNIVERSITY

M.A., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor David Kotz

The last thirty years in the US have been characterized by rising financial profits as a share of total profits and the growth of banking activities yielding non-interest income. These developments pose two questions. First, what are the social relations enabling and sustaining financial profits and what are their macroeconomic sources? Second, what do these trends imply for the nature of banking and what kind of theory of banking can capture them? This study addresses these questions and makes four contributions.

First, a Marxist theory of banking is developed to capture the transformation of banking drawing on two characteristics: first, emphasis on liquidity provision through exchange of promises to pay among credit participants and, second, explicit connection between bank revenues and macroeconomic aggregates (wages, profits, assets). It is shown that a Marxist theory of banking can be a more general theory of banking retaining strengths of other approaches and overcoming their limitations.

Second, it is shown that the corporate form of business organization and the attendant capital markets create opportunities to extract a range of profits with similar characteristics, i.e., capital gain-like revenues. The key features of their simplest form – founder’s profit – are shown also to hold for securitization revenues and, partly, for profit from mergers and acquisitions. These revenues hinge on wealth transfers across the society and, therefore, differ from profits from production.

Third, by bringing together the Marxist theory of banking and the analysis of capital gain-like revenues, liquidity provision is shown to form the basis for banks’ sharing in capital gain-like income. The core functions of banking can, therefore, co-exist with, and even form the basis, for a significant transformation of bank revenues. Examination of the multiplicity of forms of capital gain-like revenues shows that their extraction is the common driving force behind the apparently heterogeneous activities associated with the transformation of banking.

Fourth, empirical analysis of the US bank holding companies confirms that capital gain-like revenues were a significant part of bank revenues in 2001-2010. Given the rising household vulnerability toward wealth transfers, this trend suggests a reinstatement of predatory aspects of finance in contemporary capitalism.

CONTENTS

| | Page |
|---|------|
| ACKNOWLEDGMENTS | iv |
| ABSTRACT | v |
| LIST OF TABLES..... | x |
| LIST OF FIGURES..... | xi |
| CHAPTER | |
| 1. INTRODUCTION | 1 |
| 2. TRANSFORMATION OF BANKING: THEORY AND EMPIRICAL EVIDENCE. MAINSTREAM LITERATURE REVIEW | 8 |
| 2.1. Introduction..... | 8 |
| 2.2. Decline in Banking?..... | 11 |
| 2.3. Transformation of Banking..... | 16 |
| 2.4. Transformation of Banking: Where Do Profits from the “Transformed Banking” Come from? | 21 |
| 2.5. Mainstream Theories of Financial Intermediation and Banking | 28 |
| 2.6. Mainstream Theories Reconsidered: Theories of Transformation of Banking, or A Quest for a New Content of Banking..... | 39 |
| 2.7. Conclusion | 48 |
| 3. BANKING, TRANSFORMATION OF BANKING, AND BANK PROFIT IN THE THEORY OF THE MONETARY CIRCUIT..... | 52 |
| 3.1. Introduction..... | 52 |
| 3.2. Schumpeter and Keynes: Immediate Origins of the Circuitist Theory of Banking..... | 54 |
| 3.2.1. Schumpeter: Banking for Development and Creation of Purchasing Power...54 | |
| 3.2.2. Keynes: Banks and Liquidity Provision | 64 |

| | |
|---|-----|
| 3.3. Theory of the Monetary Circuit: An Approach to Banking and Its Transformation..... | 70 |
| 3.3.1. Banking in the Theory of the Monetary Circuit..... | 71 |
| 3.3.2. The Circuitist Theory of Banking in the Context of the Transformation of Banking | 76 |
| 3.4. Bank Profit in the Theory of the Monetary Circuit: Interest as Rent | 82 |
| 3.5. Financial Profit in the Literature on Financialization | 90 |
| 3.6. Conclusion | 102 |
| 4. MARXIST THEORY OF BANKING AND BANK PROFIT | 105 |
| 4.1. Introduction..... | 105 |
| 4.2. Bills View vs. Goldsmiths View..... | 107 |
| 4.3. Transformation of Banking..... | 113 |
| 4.4. Marxist Theory of Banking in the Light of Other Theories of Banking..... | 123 |
| 4.5. Bank Profit..... | 129 |
| 4.5.1. Bank Profit in Marx and Hilferding..... | 130 |
| 4.5.2. Bank Profit and Loanable Capital..... | 138 |
| 4.5.3. Bank Profit: Beyond Marx and Hilferding | 143 |
| 4.6. Conclusion | 151 |
| 5. POLITICAL ECONOMY OF CAPITAL GAINS: THROUGH THE LENS OF TWO SETS OF PRICES OF CAPITAL GOODS..... | 155 |
| 5.1. Introduction..... | 155 |
| 5.2. From the Dual Price System to a Theory of Investment: Keynes, Tobin and Minsky | 157 |
| 5.3. From the Dual Price System to the Nature of Financial Profits: Hilferding..... | 161 |
| 5.4. Reading Minsky through Hilferding’s Lens. What are Capital Gains?..... | 165 |
| 5.5. Characteristics of Capital Gains: Beyond Hilferding and Minsky | 175 |
| 5.6. Capital Gains Come in Many Forms | 179 |

| | |
|---|-----|
| 5.7. Forms of Profit due to the Gap between the Two Sets of Prices of Capital and the Transformation of Banking..... | 180 |
| 5.7.1. Profit from IPO | 181 |
| 5.7.2. Profit from Mergers and Acquisitions (M&A)..... | 196 |
| 5.7.3. Securitization, or Investment Banking for Banking | 203 |
| 5.8. Conclusion | 212 |
| 6. REVENUES AND PROFITS OF THE US BANK HOLDING COMPANIES: MULTIPLICITY OF FORMS – UNIFORMITY OF SOURCES..... | 214 |
| 6.1. Introduction..... | 214 |
| 6.2. Transformation of Banking: Illusory Profits, Mirages, and Rents | 217 |
| 6.3. Revenues and Profits of the U.S. Bank Holding Companies: An Empirical Analysis | 226 |
| 6.4. Beyond the Apparent Heterogeneity of New Banking Activities: Profits from Redistribution of Assets..... | 245 |
| 6.5. Conclusion | 258 |
| 7. CONCLUSION..... | 260 |
| 7.1. Comparative Advantage of a Marxist Theory of Banking | 261 |
| 7.2. A Range of Capital Gain-Like Revenues | 265 |
| 7.3. Extraction of Capital Gain-Like Revenues as a Driving Force behind Many Aspects of the Transformation of Banking..... | 269 |
| 7.4. Significance of the Capital Gain-Like Revenues for the US Bank Holding Companies in 2001-2010..... | 271 |
| BIBLIOGRAPHY..... | 274 |

LIST OF TABLES

| Table | Page |
|---|------|
| 1. Net change in the fair values of financial instruments accounted for under a fair value option in comparison to other capital gain-like revenues (billion of \$) and as a share of non-interest income (in percent)..... | 237 |
| 2. Capital gain-like revenues as a percentage of total non-interest income and as a percentage of total revenue, all BHC and the largest BHC, simple averages for 2001-06 and 2009-10. | 241 |
| 3. Capital gain-like revenues as a share of total non-interest income and as a share of total revenue: ratio of the largest banks to all BHC..... | 241 |
| 4. Total economy financial assets, net worth, and financial assets less liabilities as a multiple of GDP (US, 1946-2011), simple average..... | 248 |
| 5. Household financial assets, net worth, and financial assets less liabilities as a multiple of disposable personal income (US, 1946-2011), simple average..... | 253 |

LIST OF FIGURES

| Figure | Page |
|---|------|
| 1. Distribution of gains from trade in a bill of exchange | 147 |
| 2. Pre-tax financial sector profit as a share of total pre-tax domestic profit (US, 1945-2011, different sectors). | 227 |
| 3. Pre-tax financial sector profit, nonfinancial sector profit, and GDP (US, 1970-2011). Index: 1970 = 1. | 228 |
| 4. Pre-tax profits of the financial sector, commercial banks, and BHC (US, 1970-2011). Index: 1970 = 1. | 230 |
| 5. Total interest income, non-interest income, and realized capital gains (US BHC, 1986-2010). | 231 |
| 6. Noninterest income as a share of net operating revenue, total revenue, and GDP (US BHC, 1986-2010). | 232 |
| 7. Composition of non-interest income of the US BHC, 2001-2010. | 233 |
| 8. Capital gain-like revenues (US BHC, 2001-2010). | 236 |
| 9. Total capital gain-like revenues, as a share of income and profit (US BHC, 2001-2010). | 238 |
| 10. Return on equity and return on assets, US BHC, 2001-2010. | 239 |
| 11. Capital gain-like revenues as a share of total non-interest income (US BHC, 2001-2010). | 240 |
| 12. Capital gain-like revenues as a share of total revenues (US BHC, 2001-2010). | 241 |
| 13. Share of the total capital gain-like revenues of the U.S. BHC sector accruing to the largest BHC. | 242 |
| 14. Share of realized capital gains of the U.S. BHC sector accruing to the largest BHC. | 243 |
| 15. Share of founder's profit of the U.S. BHC sector accruing to the largest BHC. | 243 |
| 16. Share of securitization income of the U.S. BHC sector accruing to the largest BHC. | 244 |
| 17. Total economy financial assets, net worth, and financial assets less liabilities as a multiple of GDP (US, 1946-2011). | 248 |

| | |
|--|-----|
| 18. Households financial assets, net worth, and financial assets less liabilities as a multiple of disposable personal income (US, 1946-2011)..... | 253 |
| 19. Household non-deposit financial assets as a share of household total financial assets (US, 1945-2011). | 254 |
| 20. Composition of the household financial assets as a multiple of disposable personal income (US, 1946-2011)..... | 256 |

CHAPTER 1

INTRODUCTION

The past three decades in the US economy have been characterized by two trends. First, revenues and profits accruing to financial institutions have been rising as a share of total profits. Profits accruing to the US financial sector rose from 13 percent of total domestic profits in the 1950-60s to 32 percent in 2000-11, on average. Commercial banks profits as a share of the total doubled during the same time period – from 6 to 12 percent. Financial sector value added increased from 2 percent of GDP after the World War II to 9 percent in 2010. Second, the character of banking activities has changed too. Banks have been increasingly engaged in underwriting, mergers and acquisitions, securitization, and trade in financial assets, receiving a rising share of their revenues in the form of non-interest income.

These two developments pose a question: What do these changes mean, with respect to the nature of financial profits and the character of banking business? First, what are the sources of the forms of profit that rose to prominence in the course of financialization? There are two dimensions of this question. What are the social relations that have made a systematic extraction of these forms of profit possible and sustainable, on which grounds can they be justified? And what are the macroeconomic sources of these gains? Second, how do the new banking activities relate to the fundamental nature of banking business? Do these new activities represent a major departure from what banking is? And what kind of theory of banking is required to capture the two trends at hand?

These questions have attracted attention in the literature. The rise in financial profits and its significance have been discussed from several perspectives. Epstein and Jayadev (2005), Krippner (2005), and Reid (2008) documented a rise in financial profits as a hallmark of financialization. These profits are often seen as rents – zero-sum income transfers (Bossone 2001a, 2001b, Graziani 2003, Parguez 2004, Turner 2010). Pollin (1996) and van Treeck (2009) emphasized a necessary connection between financial profits and value created in production. Others, by contrast, stressed a fictitious character of some of the financial profits due to either accounting standards (Kerr 2011) or risk illusion (Crotty 2008, Haldane, Brennan & Madouros 2010). The fictitious character of these gains implies the “paper gains” are inevitably offset by subsequent “paper losses” of the recipients of these gains. It also implies mismeasurement of the financial sector output and GDP.

The present study contributes to this literature by identifying a range of financial profits that have similar characteristics – capital gain-like revenues. It is argued that financial profits in general come from different sources and therefore cannot be treated as a homogeneous rent, even less so when the composition of bank revenues is changing suggesting some structural shifts. At the same time, the multiplicity of the concrete forms of profits can also be misleading, and a number of apparently heterogeneous activities, such as initial public offerings and underwriting, mergers and acquisitions, securitization, and trade in financial assets, are driven by similar forms of profit. These capital gain-like revenues are shown to retain a connection to the output created in the sphere of production, but this connection is more complex than a simple redivision of the current output. The capital gain like-revenues are related to components of the expected future

output (wages or profits), but at the moment of accrual they come from a transfer of monetary wealth – either immediately (if a financial transaction is performed with the ultimate asset holder’s own funds, savings), or after some point in time (upon repayment of debt if the transaction is done with borrowed funds). The capital gain-like revenues therefore bear no relationship to and under certain conditions can exceed the current output. Although accounting practices and excessive risk taking have further boosted these gains and although some of these new banking practices have created favorable conditions for a crisis and asset devaluation, the existence of the capital gain-like revenues is not predicated upon future losses of the recipients of these gains. They exist even in “normal times” due to wealth transfers associated with differences in the required rates of return.

Second, the spread of new banking practices has brought about literature on the transformation of banking, at first empirical, subsequently followed by theoretical contributions. The changes in banking have revealed weaknesses of the information-theoretic approach that had been a dominant mainstream theory of banking since the late 1970s. A conclusion of this theory that the core banking functions were in decline sat uncomfortably with a rise in the share of total domestic profits accruing to these institutions. This suggested that providing a solution to information and agency problems, as important as they are, might not be the essence of banking, and created a theoretical vacuum calling for a new theory of banking. As a response to weakness of the dominant theoretical framework, new mainstream approaches to banking appeared since the 1990s. Instead of asserting a decline in the banking functions, these new approaches focus on an emergence of new banking functions (risk management) or new forms of performing old

functions (liquidity provision on demand on both sides of the balance sheets, asset transformation, creation of informationally-insensitive debt). As a result, these new theories are better able to explain how the transformation of banking can co-exist with a rise in bank revenues and profits, but none of them has taken the place that had belonged to the information-theoretic approach as a coherent theoretical framework. A contender among the heterodox theories was the monetary theory of production that defined banking as money creation ex nihilo and located it in the circuit of capital by focusing on lending for productive purposes.

The present study contributes to this literature by suggesting an alternative theoretical framework for understanding the nature of banking business and its transformation. A proposed framework can be seen as a more general theory, nesting some of the existing approaches as its particular moments. Grounded in the bills view and classical political economy, this theory of banking has a dual foundation. A core function of banking is understood as liquidity provision through an exchange of promises to pay, and bank revenues are explicitly connected to macroeconomic aggregates (profits, wages, assets). A relative autonomy of bank functions and bank revenues allows for different degrees of their transformation. An application of this theory to the ongoing changes in banking shows the fundamental nature of banking business has remained unchanged. The current transformation of banking represents a periodically occurring change in the concrete forms of promises to pay and mechanisms of their exchange. On the other hand, a rise of the corporate form of business organization and capital markets has resulted in a much more significant change in banking by raising to prominence profits that have a distinct macroeconomic source – monetary assets scattered across the society. This aspect

of the transformation of banking has not been captured by other theories of banking that, unlike the suggested approach, lack an explicit connection between bank profits and the key macroeconomic aggregates.

The methodology of this study involves a combination of, first, a history of economic thought and a critical analysis of theoretical paradigms and, second, an empirical investigation of the trends in revenues of the US bank holding companies. From a methodological standpoint, the present study has three stages. The first step uncovers strengths and limitations of the current mainstream and heterodox approaches to banking by using them as a lens for understanding the transformation of banking. The second step involves using insights from earlier economic thought to show strengths of an alternative (Marxist) theoretical framework to banking and bank profits that would capture the ongoing transformation of banking and overcome the limitations of other existing approaches, both mainstream and heterodox. In particular, a discussion of the bills view and goldsmiths view on banking allows for developing an approach to banking as an exchange of promises to pay. Steuart's (1770a, p. 206) concept of "profit upon alienation" is used to establish a macroeconomic source of profit distinct from profit from production. The idea of the dual price system of capital goods associated with theories of investment by Keynes, Tobin, and Minsky are used to develop an analytical tool for an analysis of capital gain-like revenues. Finally, the third stage empirically examines trends in revenues of the US bank holding companies since 1986 – first using the accounting categories as given in the income statements and then applying the concept of the capital gain-like revenues. An application of this category to an empirical work uncovers a common driving force behind many faces of the transformation of banking that would

have been overlooked based on accounting categories and shows the significance of the capital gain-like revenues in total bank revenues.

The study has the following structure. Chapter 2 overviews the mainstream empirical and theoretical literature on banking and its transformation, and concludes that the information-theoretic approach to banking has revealed its weaknesses and a resulting theoretical vacuum has been only partially filled by the new approaches to banking. Chapter 3 considers a take on banking within the monetary theory of production as an alternative to the mainstream theories. It is argued that although this theory explicitly locates banking among the other sectors and connects it to macroeconomic aggregates, it has a limited explanatory power due to a too narrow take on the functions of banking and an unnecessarily oversimplified link between banking and the circuit of capital. Drawing on a discussion on two takes on banking in the history of thought, the bills view and the goldsmiths view, Chapter 4 locates a Marxist theory of banking in the history of thought. A Marxist theory is shown to have a dual foundation – an exchange of promises to pay to provide liquidity as a core function of banking and a relationship between bank profits and flows of value. Such theory is argued to be a more general theory of banking locating other theoretical approaches as its particular instances reflecting realities of banking in a particular historical period. Chapter 5 shows that a rise in the corporate form of business organization and capital markets has created an opportunity for extraction of capital gain-like revenues taking a number of forms. Founder's profit is the simplest form of such gain which has similarities with profit from securitization and to some extent profit from mergers and acquisitions. A dual price system for capital goods which typically forms a basis for theories of investment is shown to be a helpful analytical tool for understanding

capital gain-like revenues. Due to their function of liquidity provision, banks are well suited to facilitate these transactions and on these grounds to share in the capital gain-like revenues. Thus, the transformation of bank revenues is argued to be more pronounced than the transformation of bank functions. Chapter 6 presents the results of an empirical examination of revenues accruing to the US bank holding companies. The chapter establishes that behind an apparent heterogeneity of the new banking activities lies a fundamental similarity – a turn to extraction of capital gain-like revenues – and documents that these revenues have been a significant source of bank income in 2001-2010. Household wealth is shown to be a significant source of the capital gain-like revenues, with a rise in the household financial assets in the neoliberal era indicating a rising vulnerability of households to such wealth transfers. A rise in the capital gain-like revenues in the context of a rising vulnerability of households to wealth transfers associated with these profits is argued to be a reinstatement of predatory dimensions of finance on a new foundation.

CHAPTER 2

TRANSFORMATION OF BANKING: THEORY AND EMPIRICAL EVIDENCE.

MAINSTREAM LITERATURE REVIEW

2.1. Introduction

The present chapter critically reviews the mainstream literature showing how it approaches the transformation of banking and bank profits since the 1980s. This overview will locate the present study in the already existing debate. Three main conclusions of this chapter are relevant for our analysis in the subsequent chapters. First, the mainstream empirical literature acknowledges the transformation of banking since the 1980s. Nevertheless, its empirical and theoretical account of this transformation is limited. Second, the existing debate shows the importance of looking at bank profits and revenues – for two reasons. It is the dynamics of bank profits that revealed the transformation of banking behind the apparent decline in banking that was widely believed to be taking place in the early 1990s. Furthermore, the content of this transformation can be best captured through the lens of the changing composition of revenues, and not so much from bank balance sheets. Nevertheless, the mainstream treatment of profits is limited, and the aim of the subsequent chapters is to discuss the aspects of bank profit neglected by the mainstream approaches. Third, the transformation of banking reveals the limits of the mainstream theory of banking in understanding the nature of this transformation and analyzing bank profits. Thus, the aim of the present study is to fill this gap by developing an alternative account of both the transformation of banking and bank profits.

The chapter is organized as follows. Section two discusses the empirical literature on the decline in banking which was the first wave of studies addressing the changes in banking since the 1980s. It is shown that although there were good reasons to believe that banks were declining, this early literature showed that the changes should rather be understood as a transformation of banking, not its decline. This discussion also put the role of profits to the forefront. The third part reviews the empirical literature on the transformation of banking, summarizing the changes in banks assets, liabilities, and off-balance-sheet activities as depicted in the mainstream literature. It also discusses limitations of this literature, with an emphasis on its empiricist bias and absence of macroeconomic perspective. Section four reviews the mainstream discussion of bank profits in the course of the transformation of banking. The section summarizes the main themes in this literature (such as the rise in fee income, bank revenues and diversification benefits, the link between the structure of revenues and bank size, and the discussions on particular types of non-interest income), and shows their limitations. The present study aims at overcoming these limitations by discussing aspects of profit neglected by the mainstream approaches.

The second half of the chapter addresses the mainstream theoretical literature showing how this empirically observed transformation of banking fits with the dominant theory of banking. Section five reviews the mainstream theories of banking – its first wave associated with the portfolio, or asset transformation theories, its second wave represented by the information-theoretic approach, and the relationship between the two. It discusses features of these theories preventing them from being a good foundation for understanding the transformation of banking. It is also shown that the information-

theoretic approach, dominant at the time when the transformation of banking is discussed, provides a limited treatment of this transformation due to focusing on the quantitative shift in banking and failing to see qualitative changes. In this sense the mainstream theoretical literature lags behind its empirical stream. Furthermore, the quantitative decline in banking predicted by the second wave theories of banking cannot be reconciled with a rising bank profit share, which calls for a new theoretical explanation of banking.

New theoretical conceptualizations of banking that emerged as a response to the inability of the information-theoretic approach to address the transformation of banking are discussed in part six of the chapter. Among them are approaches centered on risk management, liquidity provision on both sides of banks' balance sheets, creation of informationally-insensitive debt, and credit intermediation through asset transformation. The section discusses the core ideas of these theories indicating their strengths compared to the previous theories of banking. It is argued that, in spite of being more adequate to understanding the transformation of banking, these new approaches are still based on a banking view because they do not provide a theoretical foundation for understanding certain dimensions of the transformation of banking. Furthermore, these theories are characterized by dichotomous thinking about this transformation, because they either deny it as a qualitative change or show that the content of banking has undergone a major change. As a result, there remains a need for a new theory of banking that would overcome these limitations.

Section seven concludes by summarizing the main points of the chapter, including the ideas not addressed in the mainstream empirical and theoretical literature, and by

formulating the requirements for an alternative theory of banking capable of capturing the transformation of banks since the 1980s.

2.2. Decline in Banking?

In the early 1990s there has emerged vast literature addressing the question whether banks have been dying out (Edwards 1993, Boyd, Gertler 1994, and an update on them by Feldman, Lueck 2007, Kaufman, Mote 1994, with a continuation of the discussion in Schmidt 1999, Herring, Santomero 2000, Samolyk 2004). It was a natural response to difficulties in the US banking sector in the late 1980s and to an apparent tendency of banks to lose market share to other financial intermediaries. Banks were facing severe competition on both sides of their balance sheets.

On the asset side, banks were losing importance in credit provision to non-financial corporation that were turning to raising funds directly through markets. This process was facilitated by emergence of junk bonds, finance companies, and commercial paper market. Banks were also losing market share in consumer lending which in 1990-2006 dropped from 50 to 30 percent of total consumer loans outstanding. It was a consequence of a rise in lending by non-bank financial intermediaries (especially finance companies and government-sponsored enterprises), by non-financial intermediaries (such as General Motors, for example), and a spread of securitization, which resulted in 30 percent of consumer loans being securitized in 2006¹ (Feldman, Lueck 2007, p. 42, Herring, Santomero 2000, p. 30, Samolyk 2004, p. 48, 51).

¹ Feldman and Lueck (2007, p. 42) notice that securitization represents both a competitive threat and a benefit for banking.

Banks were losing market share on the liability side of their business, as well. The usual argument in the literature is that due to prolonged inflation, among the other reasons, households have become more sensitive to interest rates differentials. A demand for deposits with higher rates of return was met by proliferation of money market mutual funds (MMMMF).

These processes of banks losing importance in providing what used to be their traditional services led to emergence of the literature which essentially argues that, in spite of all this evidence, the conclusion about banks' dying out is premature. Although there are some differences among authors, there have been several lines of reasoning common to all of them.

First, even though banks' assets have been declining as a share of assets of other financial intermediaries, they have been rising as a share of GDP. So if banks are losing importance, it is so only with respect to other financial institutions, but not with respect to the economy as a whole.

Second, it was noticed that off-balance sheet activities have been playing an increasingly important role in banking business. Once banks' assets are adjusted for these practices through different ways of imputing on-balance-sheet equivalents of the off-balance sheet activities, this new measure of bank assets has not been declining and remains at an average 45 percent of total financial sector assets since the mid 1980s (Boyd, Gertler 1994, Feldman, Lueck 2007, p. 43-44)².

This observation was pushed to its logical limit by arguing, thirdly, that using assets as a measure of banks market share is not only quantitatively misleading due to the

² Another adjustment usually made is for foreign banks entry. Once foreign banks' assets are taken into account, it further mitigates the apparent decline in banks' market share.

need for the above mentioned adjustments, but also conceptually incorrect. Assets are an inadequate measure of banks' market share for two reasons (Kaufman, Mote 1994, p. 4). First, there is no one-to-one correspondence between services provided and value of the assets, hence, assets do not reflect the importance of banking functions. Second, assets are a stock category, whereas output, including bank output, should be measured by a flow category. On these grounds it was concluded that banks' share in total profits would be a better measure of their market share. Some authors made it the core argument of their study (Samolyk 2004). A few empirical analyses that used profit share as a more adequate reflection of banks market share found that it has been stable or rising, implying no decline in the relevance of banks, compared to other financial intermediaries (Feldman, Lueck 2007, p. 47, Kaufman, Mote 1994, p. 4, Samolyk 2004, Mishkin 2001, p. 272).

Fourthly, a related attempt to use a different measure of banks' market share was based on value added. Similar to profit, it was found to fluctuate around 50 percent of value added by all financial intermediaries, with no secular decline (Feldman, Lueck 2007, p. 46).

Finally, the other indication of banks' not dying out is new entries into the banking business, which signifies that the industry is viewed as potentially profitable (Feldman, Lueck 2007, p. 47).

The relevance of the early 1990s discussion on the decline in banking is in its drawing attention to the fact that banks have not been dying out, they have rather been evolving. This shifted the focus of analysis from a decline in to a transformation of banking. Put differently, it revealed that the actual concern of theoretical and empirical

studies should be not a decline in banking. For this apparent decline is both a cause of the transformation of banking and a form in which this transformation appears. It is a cause, because it is the competitive pressure on banking firms and changing overall macroeconomic conditions in which they operate that made banks “find other activities which are profitable and transform themselves into viable entities which compete with other firms called ‘non-banks’” (Gorton 1994, p. 116). It is a form, because one of the major constituents of this transformation is proliferation of off-balance-sheet activities that by their nature do not appear on bank balance sheets, creating an illusion of a decline in banking, if banks’ market share is measured by assets. In this sense, the literature on the decline in banking set a context for another stream of the literature that emerged in the early 1990s – that on the transformation of banking.

In addition to this key insight, there are a few other important contributions of that discussion to the way banking is understood in empirical and theoretical literature.

First, it put this “decline” in banking in a broader historical context by showing that, as is often the case with economic processes, there is nothing new either about the apparent decline in banking, or about the discussions about it. Already since the 1860s, there has been a secular trend of a decline in bank assets and loans as a share of those of the financial system as a whole, with this process being put on hold some times and reversed other times (Edwards 1993, p. 7, 16-17, Edwards, Mishkin 1995). A period resembling the 1980s the most were the 1920s, when a decline in the bank asset share

was especially pronounced (Herring, Santomero 2000, p. 27) and investment banking was on an upsurge (Kaufman, Mote 1994, p. 7, data from Goldsmith)³.

Secondly, the fact that a few studies concluded on inadequacy of measuring banks market share by assets is an example of the empirical literature catching up with the theories of financial intermediation as they evolved in the 1980s. As was mentioned above, Kaufman and Mote observed that assets are an inadequate measure of banks market share due to their not being related to the volume of services provided and due to their being a stock category. But this effectively means that empirical studies measuring the bank market share by assets implicitly correspond to a “productionist” view on bank services, with banks producing loans. It essentially reflects the empirical literature lagging behind the mainstream theoretical developments that already by the 1980s went far beyond the perspective on banks as producers of loans and shifted the focus of analysis to banks as specialists producing information and creating incentives. Thus, Kaufman and Mote’s suggestion to use the profit share, which was stressed a decade later by Samolyk (2004), is an important step in the empirical literature coming to terms with changes in the mainstream theories.

A corollary to this is that the discussion on a decline in banking effectively put the question of banking profit to the forefront and set the conditions for its analysis as a key indicator. Nevertheless, the empirical literature correctly emphasizing the role of the bank profit as a more adequate measure of the market share does it in a limited way. It relates bank profit to that of financial intermediaries, and concludes that banks receive a constant

³ The idea of “retrogression” of monetary intermediaries relative to financial intermediaries in general – what is now called disintermediation – going back to the early 20th century can also be found in theoretical literature (Gurley, Shaw 1956, p. 260, 269).

share of profits. But from a macroeconomic and political economy perspective, a more important indicator of the role of banking in the economy as a whole is a share of bank profits in total domestic profits. As was shown in the previous chapter, this profit share has increased over the past several decades. It also seems to move independently of the total domestic profit. By neglecting this aspect, the mainstream literature cannot pose a theoretical question of the content of bank profit, as compared to that of non-financial firms. Thus, the emphasis on the bank profit share is important, but it needs to be pushed even further. By the same token, the question of profit also arises in the context of the transformation of banking. As it will be discussed in the third section of the present chapter, paradoxically, in spite of acknowledging the role of revenues for understanding banks market share, profit as a lens for an analysis of the transformation of banking was not widely used, and there are objective reasons why it could not happen.

Finally, although the literature on the decline in banking was an example of a reconciliation of the mainstream empirical and theoretical studies, at the same time the transformation of banking which became obvious in the course of the discussion posed challenges to the theoretical literature threatening the newly emerged reconciliation. We will come back to this issue in the subsequent parts of the present chapter.

2.3. Transformation of Banking

As was argued above, the major contribution of the discussion on the “decline” in banking is that it set a general context for shifting the focus of attention away from the decline per se to the transformation of banking⁴. This issue was picked up by several

⁴ This transformation of banking is usually viewed as a part of a broader process – that of a transformation of credit markets, or financial systems in general.

empirical studies (Samolyk 2004, Edwards, Mishkin 1995, Boyd, Gertler 1993, DeYoung, Rice 2004a).

The shift in the focus of analysis from a decline in to a transformation of banking effectively signified a shift in the conceptual framework underlying the empirical studies – from an institutional to a functional perspective on banking. This was accompanied by a shift of the vantage point for empirical work – from an analysis of bank market share in total assets and its different forms (loans extended to businesses, households, etc) to an analysis of composition of banks balance sheets and changes thereof. The transformation of banking analyzed through this lens is usually described along the following lines.

On the asset side, the literature stresses three major changes. First, bank loans as a share of assets has risen due to a decline in reserve requirements, development of money markets causing a decline in precautionary holding of securities, and a rise in liquidity of certain types of loans due to securitization (Boyd, Gertler 1993, p. 324, DeYoung, Rice 2004a, p. 40)⁵. Second, commercial and industrial lending (C&I lending, from now on) has declined as a share of total assets due to firms increasingly funding themselves via commercial paper, junk bonds, and finance companies (Boyd, Gertler 1993, p. 324), as well as due to a later on rise in hedge funds, venture capital, and private equity funds providing the same set of functions for non-financial corporations (Feldman, Lueck 2007, p. 9). This decline in C&I lending has been especially pronounced for large banks (DeYoung, Rice 2004a, p. 41). Third, mortgage lending, especially commercial mortgage lending, has been rising as a share of bank assets (Boyd, Gertler 1993, p. 325).

⁵ From a theoretical perspective it implies a certain degree of substitution between liquidity of the asset side of the balance sheets and ease of raising liquidity on the liability side. It is similar to Hicks' point on substitutability between liquid assets and borrowing power (one's ability to increase liabilities, or what he calls "invisible assets") (Hicks 1989, p. 66).

On the liability side, the literature on the transformation of banking usually points out a decline in the role of deposits – the process referred to as disintermediation. The opposite side of the same coin is banks’ turning to the money market – to purchased money (large time deposits, federal funds, and repo), which raises their exposure to liquidity risk. Rising leverage is the other change on the liability side.

In addition to these changes on the asset and liability sides of the balance sheets, the other aspect of the transformation of banking is a rise in their off-balance-sheet activities. The three distinct types of these activities emphasized in the literature are loan sales and securitization, standby letters of credit and loan commitments, and derivatives. The discussion of the off-balance-sheet activities even in the early literature is usually accompanied by a caveat that they do not always and not necessarily represent new lines of business: “In many cases, further, the growth of off-balance sheet activities reflects only superficial rather than substantive changes in the nature of banking” (Boyd, Gertler 1994, p. 8).

In spite of the major advantage of this literature in shifting the object of concern away from a decline in banking to its transformation, there are two major limitations of these empirical studies.

Firstly, and most importantly, an analysis of the transformation on the asset side of the balance sheets – changing composition of lending – has a strong empiricist bent. The taxonomy of assets is driven by the empirically given types of loans, and is not informed by an understanding of conceptual differences among different types of lending. Most of the studies stress the rise in loans secured by real estate, without a further distinction between mortgages to households and to nonfinancial corporations. This

distinction is nevertheless crucial, as the former has a different economic status from the usual bank loans to businesses, whereas the latter in many cases can be a formal substitute for these loans. When it comes to the mainstream empirical studies, lending to households appears as a separate category only when the empirical data make it so, as it is the case with loans to individuals consisting of consumer loans and credit card loans.

The only exception is a study by Boyd and Gertler (1993) who mention that a rise in lending secured by real estate comprises heterogeneous loans. But when it comes to an empirical analysis of those, the authors relate different types of mortgage lending to total mortgages, instead of total assets or loans, and conclude that commercial mortgage lending has become more important – a conclusion opposite to the one we reached in the previous chapter based on an examination of the empirical data. The reasons behind the difference between our conclusion and that by Boyd and Gertler lies in the fact that in the period studied by them home mortgages were declining as a share of total mortgages. Nevertheless, a rapid rise in lending secured by real estate as a share of total loans overpowers this decline, so that home mortgages were increasing as a share of bank assets – a conclusion one cannot reach looking at the share of home mortgages in total mortgages. In spite of overlooking it, Boyd and Gertler deserve a credit for, first, at least stressing the difference between the two types of mortgages, and second, for also mentioning that this decline in home mortgages as a share of total mortgages in the banking sector does not correspond to the general trend in mortgage lending in the economy as a whole, where home mortgages represent a constant 80 percent fraction of total mortgages (Boyd, Gertler 1993, p. 325)⁶. Thus, they treat it as a difference between

⁶ Appreciation of the role of lending to households, although in a different context, can also be found in Samolyk who, discussing consumer credit, stresses the difference between funding loans and originating

banks and other financial intermediaries, instead of seeing that banks are a part of the general process of rising home mortgage lending.

The lack of appreciation of the conceptual difference among different types of lending within the mainstream literature is obvious in a study by DeYoung and Rice (2004). In the summary tables on changes in composition of bank assets, they have data on the ratio of residential mortgages to loans, showing a rise in the size-weighted averages of the ratio and a constant ratio for the unweighted average (DeYoung, Rice 2004a, p. 41). It implies a rise in household mortgage lending by large banks (that dominate the size-weighted averages) and its constant share for “typical” banks. Commercial mortgages exhibit the opposite tendency being a constant share of total loans for large banks and rising only for “average” banks. In spite of having these data, DeYoung and Rice do not discuss these trends and their implications. Instead, in the discussion of the transformation of banking one finds a usual argument about a rise in lending secured by real estate, without further specification.

A second limitation of the empirical studies on the transformation of banking through the lens of the systematic changes in their activities as reflected in the structure of their balance sheets is that there are no recent studies of this issue, except for those looking at individual dimensions of this process. There was some interest in the question in the mid 1990s, and a few updates in the early 2000s. Since then, there have been studies of individual dimensions of the transformation of banking, not the process as a whole. But what can one say about the transformation of banking as a systematic phenomenon since the early 2000s? Did it stop? Is it reduced to a few tendencies with no

them. It allows her to conclude that although the role of banks in the former has been declining, it has been rising in the latter (Samolyk 2004, p. 48-49, 51).

unifying principle behind them? Was there a reversal of those tendencies, or a continuation of the same trends? Has the transformation of banking acquired new features? Mainstream literature does not give answers to these questions.

Summing up, the mainstream empirical literature does acknowledge a structural transformation of banking since the 1980s. Nevertheless, it does it in a limited, one-sided way. For this reason Erturk is not quite correct arguing that “both the mainstream economics and the political economy literatures on banking do not seem to have noticed the transformation of the banks’ balance sheets and their revenue sources since the early 1990s. Such data on banking show that banking has become increasingly a business where their role in providing finance to the productive sector is no longer the dominant one. Financing households, selling investment products to households and transacting between themselves through instruments of financial innovation have become the dominant activity of financial intermediation” (Ertürk 2009, p. 8). The mainstream empirical literature did notice the transformation of banking. Moreover, what Erturk calls banks’ “transacting between themselves” has always been an integral part of banking business in the form of money market transactions. So the real questions are how the mainstream literature interprets the content and implications of this transformation, and what the difference (if any) is between the traditional and the new forms of banks transacting among themselves.

2.4. Transformation of Banking: Where Do Profits from the “Transformed Banking” Come from?

It was argued above that the literature on a decline in banking put the question of profit as a proper measure of banks market share and of their role in the economy to the

forefront. As mentioned by Kaufman and Mote and turned into the key argument of a study by Samolyk (2004), the conclusion about an evolution – and not a decline – of banking becomes obvious, once one looks at bank profits, not their assets. In spite of that, the discussion of the transformation of banking has been mostly centered on the question of changes in the structure of bank balance sheets, with very little attention given to sources of bank revenues and changes in their composition. And yet that would have been the natural development of the empirical literature.

This was also noticed by DeYoung and Rice, who were effectively the first to pose the question of changes in sources of bank revenues in the course of the transformation of banking. They argue that the neglect of the income statements and composition of revenues can partly be explained by the perseverance of the traditional view of banks as financial intermediaries taking deposits and making loans (DeYoung, Rice 2004a, p. 36). Given that their traditional function is best reflected on bank balance sheets, it shapes the focus of empirical studies. Nevertheless, the transformation of banking consists precisely in changing banks' relevance as traditional financial intermediaries, hence, balance sheets become increasingly inadequate for understanding what banks do. This calls for a closer examination of banks income statements. Put differently, if one is to capture the content of the transformation of banking, one should analyze sources of banks' revenues and profits. This argument by DeYoung and Rice effectively means that revenues and profits are an adequate measure not only for showing that banks have not been dying out, but also for a study of their transformation. The question of bank revenues and profits was addressed by DeYoung and Rice and it will be the main object of the present study.

There are several themes discussed in the mainstream empirical literature on bank revenues and profits. The main argument is usually linked to a rise in non-interest (fee) income of the US commercial banks, which increased from its record low 7 percent of total revenue in 1981 to 36 percent in 2003.⁷ It subsequently declined to 26 percent of total revenue in 2007, only to bounce back to 32 percent in 2011. Commercial bank non-interest income as a share of net interest income also rose – from the average 22 percent in the late 1970s to around 75 percent in 1999-2006. There are a few relevant caveats stressed in the literature. For instance, not all the non-interest income comes from non-traditional activities. A rise in fee income reflects both a rise in non-traditional activities and a change in the mode of performing traditional banking functions, as is the case with securitization and payment services (DeYoung, Rice 2004a)⁸. Moreover, fee income is not a new phenomenon, nevertheless its trans-historical importance is sometimes obscured by movements in interest rates. For example, a rise in interest rates after World War II resulted in a decline in fee income as a share of banks operating revenue over a few after-war decades (Kaufman, Mote 1994, p. 14), whereas fee income as a share of net interest revenue remained relatively constant.

The second theme discussed is bank revenues in their relationship to diversification benefits and risk-adjusted returns. Contrary to the early literature that anticipated a reduction in volatility of bank income due to a reduction in interest rate risk and credit risk and due to diversification benefits, later studies usually conclude that the

⁷ Calculations by author based on data from FDIC, Historical Statistics on Banking, Table CB04.

⁸ The fact that fee income on credit lines is approximately equal to net interest income on a commercial loan of the same size (Boyd, Gertler 1994) can be seen as a suggestion that fee income is just a new form of revenues with the same content. Similarly, a large share of fee income comes from payment services, according to DeYoung and Rice, and Radecki (DeYoung, Rice 2004a, p. 47, Radecki 1999).

rise in non-interest income is associated with a rise in profits, but also their rising volatility, which on the net brings about a decline in risk-adjusted profits (DeYoung, Rice 2004a, Stiroh 2004). Thus, in spite of individual banks benefiting from non-interest income, on average, there are no diversification benefits.

The third issue is the relationship between non-interest income and bank size. It is argued that the share of fee income in net revenues and specific types of fee income vary with the bank size. Fee income is predominantly a large bank phenomenon. Larger banks have larger non-interest income, which comes mostly from investment banking, asset management, securitization, and insurance activities, contrary to smaller banks generating smaller non-interest income, mostly through fees on deposits (DeYoung, Rice 2004a, p. 47, Waldrop 2002).

Finally, there is empirical literature focusing on particular types of non-interest income and their specificities. For instance, it is shown that trading revenue is the most volatile component of fee income (DeYoung, Rice 2004a) and that payment driven revenues is a broader category than fees on deposit accounts (Radecki 1999). A relevant insight is that securitization income has been the main source of non-traditional bank output and the single largest contributor to the growth of fees and commissions⁹ – not even trading revenues that get most of attention (Inklaar, Wang 2007).

In spite of this attention to bank revenues and profits along the lines discussed, which is a major step forward compared to an analysis based on balance sheets only, there are several problems with and limitations of the mainstream empirical studies.

⁹ In 1990-2004, securitization revenue measured as a sum of net servicing fees and net securitization income rose from 1.2 to 11.4 percent of total bank output (Inklaar, Wang 2007, p. 24, 33, see also p. 40, 47).

Firstly, they do not use all the analytical pegs provided by the empirical studies of bank balance sheets. Specifically, there is no examination of interest revenue associated with different types of loans and there is no analysis of the relationship between the mode of bank funding and its revenues. In this sense there is some disjuncture within the mainstream empirical literature, with its different streams not fully utilizing each others insights.

Secondly, and more important, the mainstream empirical literature does not pose the question of sources of different types of revenue. Instead, it is concerned with the issues of their stability, their relationship with bank size, and so on. It is not surprising, given that the question of the origin of profit is a concern of classical political economy and Marxism, in particular. The question is usually not posed by neoclassical theories. The distinct features of those do not allow these theories to recognize the problem, let alone solve it. For this reason, the present study will fill the gap in the literature by explicitly addressing the question of the sources of financial profits, and will do it from a political economy perspective.

A corollary to the neglect of the question of the origin of profit is a misleading classification of types of banking revenues. The literature often distinguishes between traditional and non-traditional banking services. It is argued that there is no one-to-one correspondence between traditional and non-traditional services, on the one hand, and interest and fee income, on the other hand. Put differently, some fee income comes from traditional banking services.

This valid argument nevertheless misses the point that the real question is not so much whether income comes from traditional or non-traditional services, but rather the

origin of this revenue. Traditional activities are a heterogeneous category, so that different components of associated revenues can have different sources. For instance, it will be shown in the subsequent chapters that even traditional interest income can have different sources, depending on the borrower type, and although securitization is a new method of traditional lending activity, securitization income has an entirely different content from simple interest on loans. By the same token, not everything that looks like traditional payment services is actually payment services, hence, has the same source of profit. For example, revenues from securities handling, and forgone interest on credit cards imputed as securitization income from credit card receivables, considered by Radecki (1999) to be payment driven revenues, should have a separate analytical treatment. Thus, in an analysis of bank income statements, the distinction should be drawn not so much between fee vs interest income, and not even between traditional vs non-traditional banking services, but rather between bank profits in relationship to flows of value and their constituents. Hence, functional approach to bank revenues should indeed be at the heart of distinction among different types of activities and corresponding sources of revenue, but it should not be taken formally to generate envelope categories encompassing types of revenues having only formal similarities. But to accomplish it, the functional approach to banking itself needs to be modified.

A third problem with the mainstream empirical literature is its inability to see that some limitations of the mainstream theories are deeply rooted, so that it is not accidental that certain types of bank revenues cannot be discussed within the dominant theoretical framework. For example, Radecki (1999) uses his empirical study of payment driven revenues to substantiate his theoretical proposition that an emphasis on intermediation

characteristic of the mainstream theories of banking makes the theory blind to other essential functions of banks, specifically, the role of payment services. He calls for a theory taking an explicit account of it. This is a relevant insight, from our perspective. Nevertheless, it fails to acknowledge that it is not accidental that payment services are not given sufficient attention in the mainstream theories. How can a theory having information asymmetries as its core explain the necessity of payment services? A theory capable of doing it needs to have an explicit account of needs of circuit of capital and revenue¹⁰.

To conclude, the mainstream empirical literature acknowledges the transformation of banking since the early 1980s. Nevertheless, the empirical treatment of this transformation even through the lens of changes on the balance sheets is limited. For example, there is no treatment of lending to households and to non-financial corporations as distinct types of lending, even if they are forms of what seems to be a uniform category of mortgage lending. Even worse, there is a lack of attention to sources of profit in light of this transformation, which is noticed by DeYoung and Rice (2004) who initiated a discussion aiming to fill this gap. But even these attempts do not use all the analytical pegs available from the studies of the transformation of the bank balance sheets. Moreover, while focusing on a specific subset of important issues related to bank revenues, these studies neglect the question of the origin of these revenues, which is not surprising given the nature of the theoretical tradition implicitly underlying these studies. Thus, the mainstream empirical literature has a one-sided take on the transformation of

¹⁰ It will be shown below that a difference between the requirements of the circuit of capital and requirements of circuit of loanable capital itself determines the difference between revenues from dealing in money stemming from surplus value and revenues representing founder's profit having a semblance to payment services.

banking through the lens of the bank balance sheets, and a disproportional emphasis on it, with largely missing analysis of the bank income statements.

In the context of this discussion of the transformation of banking another set of issues arises. Once the mainstream empirical literature acknowledges the transformation of banking, this poses a question about the content of this transformation at a theoretical level. What do these changes mean for the nature of banking and sources of its revenue? Following Merton and Bodie functional perspective on banking, a question arises: Is it a change in the function of banking (of what banks are and what they do, hence, of social grounds for their remuneration), or a new form of the same content of activities? Is it a change in functions, or in the institutional form? Is “transformed banking” still banking? For example, Samolyk (2004, p. 50) argues that “new services provided by banks – whether the selling of mutual-fund shares to investors or the origination, sale, and servicing of loans funded by securitizations – are merely banking in different forms”. Is this approach valid? If it is still banking, what makes it so?

Before answering this question and locating the transformation of banking, we need to understand what the function of banking is, from a mainstream perspective. Hence, we need to consider mainstream theories of financial intermediation and banking.

2.5. Mainstream Theories of Financial Intermediation and Banking

The mainstream general equilibrium theory with money neutrality used to have no room for money and financial intermediation. Only in the mid 1950s, the first wave of mainstream literature emerged posing a question why financial intermediaries exist. The discussion was initiated by Gurley and Shaw, and subsequently picked up by Tobin and

Fama. These early theories conceptualize financial intermediaries as portfolio managers (Fama 1980). Financial intermediaries buy primary securities and substitute them by indirect securities with a different set of characteristics (Gurley, Shaw 1956, p. 259), which is the same as Tobin's argument about financial intermediaries matching portfolio preferences of lenders and borrowers by transforming the nature of obligations of the borrowers and assets of the lenders (Tobin 1963, p. 3-4). This service is a foundation for their receiving profit in the form of interest spread – a sort of portfolio management fee (Fama 1980, p. 46, Gurley, Shaw 1956, p. 259).

The portfolio management approach to financial intermediation pivots on the idea that there is no sharp line between money and financial assets. The boundary between the two changes historically, and in each particular case it is also contingent on specific circumstances. There is a certain degree of substitutability between money and financial assets, especially when it comes to the function of money as a store of wealth. Hence, the fact that banks' liabilities have a monetary nature cannot be used as a reason for a special analytical treatment of banks as compared to other types of financial intermediaries. As Tobin put it, the difference between the two is "of degree, not of kind" (Tobin 1963, p. 10). For this reason, for the portfolio management theorists the difference between banks and other financial intermediaries lies in the credit creation by the former being restricted not only by a profit imperative, as in the case of all financial intermediaries, but also by government regulation in the form of reserve requirements and interest rate ceilings.

These theories of financial intermediation have not been developed any further. A few decades of a relatively low interest in the issue and Friedman's shift of focus to money supply were followed by a second wave of theories of financial intermediation

that emerged in the late 1970s-early 1980s, under an umbrella of an information-theoretic approach to financial intermediation. This approach pivots on information asymmetry in different forms, with different strands within it addressing a particular form of information problems. The uniform theme is that banks are specialists in handling information asymmetry, mitigating market imperfections. If markets were perfect and did not fail, there would have been no need for financial intermediaries. But given that information is asymmetric and there exist transaction costs, there is a need for financial intermediaries.¹¹ Broadly speaking, there are two types of information asymmetry – regarding the borrowers and the lenders. Thus, financial intermediaries address information asymmetry related to both sides of their balance sheets.

According to the asset side argument, financial intermediaries emerge to produce information regarding the borrowing entrepreneur and to solve her incentive problems. There are two types of approaches here. First, there exists *ex post* information asymmetry regarding realization of returns under conditions of costly monitoring (Diamond 1984). Financial intermediaries emerge as an optimal response, helping avoid a free-rider problem and duplication of monitoring costs. They are delegated monitors that have a net cost advantage due to diversified lending. This function accounts for financial intermediaries being highly leveraged, as debt with a fixed return is an optimal contract giving a bank an incentive to monitor and resolving the problem for the ultimate lender

¹¹ This insight has broader implications for understanding the nature of the neoclassical theory. The information-theoretic approach effectively implies banks are a real-world violation of one of the assumptions of the neoclassical theory, namely, a perfectly competitive market system. By contrast, a Marxist theory, as it will become clear in chapter 4, studies banks as they are, starting with their relation to the circuit of capital. It can be argued, therefore, a Marxist theory is an aid to understanding capitalist reality, whereas the neoclassical theory of a competitive market system is an obstacle to such understanding, with the result that such an important aspect of capitalist reality – banking – can only be explained in relation to the neoclassical theory by means of violations of the core theory. I am indebted to David Kotz who has pointed out this broader implication to me.

whose returns are now guaranteed. The other interpretation of financial intermediaries handling borrowers-related information asymmetries is based on *ex ante* asymmetry between issuers of securities and their potential buyers (Leland, Pyle 1977). Moral hazard prevents direct transfer of information regarding the quality of an investment project, so there is a need for an indirect information transfer, in the form of signaling. Financial intermediaries emerge as specialists in gathering information. They hold corresponding assets, and by doing that they signal the quality of assets to other potential investors. The profits they receive in the form of interest spread represent returns to information they gathered. In both cases, information asymmetries are the primary reason why financial intermediaries exist.

There is another line of reasoning pivoting on information asymmetries related to the liability side of banks' balance sheets. Investors have unexpected liquidity demand, which is unobservable and non-verifiable. As a result, enforceable contracts cannot be written and banks emerge as providers of liquidity insurance in the form of demand deposits (Diamond, Dybvig 1983)¹². Thus, the root cause of existence of banks is information asymmetry regarding characteristics of agents and their liquidity needs. This idea is further developed by Myers and Rajan who argue that financial intermediaries remain crucial in liquidity provision even when there are capital markets as an alternative

¹² Diamond and Dybvig acknowledge that there is nothing specific to banks that would make them the sole firm capable of this type of liquidity transformation – it would hold for any firm with maturity mismatch. In that sense it is not a theory of banking, but of any firm with illiquid assets and liquid liabilities. Nevertheless, it does not pose an analytical problem for Diamond and Dybvig who retreat to empirical observations and argue that ultimately this is a theory of banks, for in reality banks bear most of the liquidity risk in an economy (Diamond, Dybvig 1983, p. 417-418). They do not see that the real question here is why it is the banks that bear most of the liquidity risk. Thus Diamond and Dybvig use as a proof what itself needs to be explained. Do banks just happen to bear this risk? Or there is something structural about this relationship? The question is addressed much later, when Diamond and Rajan show desirability of a fragile capital structure for a bank (Diamond, Rajan 2001a). It can discipline a bank but cannot discipline entrepreneurs because the former is not a creator of value, it merely transfers it.

source of liquidity, because asymmetric information causes limited participation of agents in the markets, so that banks need to fill liquidity gap and cross-subsidize (Myers, Rajan 1998). The liquidity gap emerges because not all the agents participate in the market, pushing demand for liquidity above its supply and raising the price of liquidity. By diverting a part of the liquidity demand, banks fill the liquidity gap and enhance liquidity provision by markets by raising asset prices (lowering price of liquidity). By cross subsidization Myers and Rajan mean a transfer of funds among those who do not participate in the market – a traditional liquidity insurance function of banks. Notice that the reason behind the limited participation is again information asymmetry and costs associated with acquiring information. The liquidity insurance approach is related to the “transaction accounts” take on banking (Fama 1980, Corrigan 1982, Corrigan 2000, Miller 1998), based on the idea that given that the asset side does not make financial intermediaries special, it is their liability side that distinguishes them from other firms. Their liabilities “are payable on demand at par and... are readily transferrable to third parties” (Corrigan 1982)¹³.

¹³ This primacy of the liability side gives rise to one of the theories of origin of banking and a corresponding approach to endogeneity of bank balance sheets with respect to its functions. Banks emerge as money changers. To facilitate payments, they start taking deposits. If they were to hold all the assets in a liquid form, they would have a high transformation risk – risk of a sudden change in the form of assets – which would lower their credibility. To be able to attract depositors and to make liquidity provision on demand credible, they need to hold illiquid assets to lower the transformation risk, which makes them extend loans. Thus, whereas in general holding more liquid assets makes it easier for institutions and individuals to raise capital, in the case of banks it is lowering liquidity of assets that makes it easier to attract deposits. This happens because holding a long term loan signifies a bank commitment and increases trust in bank operations. This difference between benefits of liquidity of assets for banks and non-banks constitutes the so-called “paradox of liquidity” (Myers, Rajan 1998, p. 760-761, Rajan 1996). According to this theory of banking, banks’ money-dealing endogenously turns them into deposit takers, which in turn determines the structure of the asset side of their balance sheets. Thus, the nature of bank liabilities determines the nature of its assets. Notice, it is effectively a goldsmiths view of origin of banking. Incidentally, this approach is related to Marx’s argument in volume 3 of *Capital* arguing that money-dealing capital becomes subsumed under the credit system. For an opposite approach to banks balance sheets endogeneity (assets determining liabilities), see Goodhart and various heterodox theories (Schumpeter, post-Keynesian theories, circuitists, Bossone).

It is worth discussing the relationship between the two waves of the mainstream theories of financial intermediation. This relationship is usually described as the information-theoretic approach incorporating the main insights of the portfolio theories, so that the new theory is broader and more complete. For example, in his literature review on real/financial interaction Gertler (1988, p. 574) argues that Gurley and Shaw's insights were subsumed under the information-theoretic approach which differs from the earlier theories in the methodology used. From our perspective, this is an inadequate characteristic of the relationship between the two waves of the mainstream theories. The information-theoretic approach incorporates the asset transformation aspect of banking activities only partially and, what is more important, formally. In spite of both theories being ahistorical and not being capable of distinguishing between capitalist and non-capitalist banking, the portfolio theories are at least broad enough to be able to incorporate a variety of explanations why assets are transformed. Even though an implicit assumption of this theory is differences in subjective preferences of individuals requiring this transformation, in principle there is no need to explain asset transformation by subjective differences between borrowers and lenders. Put differently, the general framework of asset transformation can be filled with different content. By contrast, the information-theoretic approach reduces all banking activities, including its asset transformation dimension, to tackling information asymmetries. This is a very specific explanation which cannot be filled with different content.

In this sense, by shifting the focus of analysis to overcoming market imperfections, the information-theoretic approach impoverishes the portfolio management theories, instead of enriching them with a new methodology. This

impoverishment is twofold. First, the information-theoretic approach reduces the content of asset transformation to a mere by-product of other banking activities related to mitigating information asymmetries. Asset transformation ceases to be an activity banks engage in as an end on its own or for reasons unrelated to information asymmetries. Secondly, even if there were no information asymmetries, there could still be a need for asset transformation, hence, for banking¹⁴. But the opposite is not true: if there were no need for asset transformation, there would be no need for financial intermediaries, hence, the information gathering and processing dimensions of banking activities would have been undertaken by direct lenders. Hence, the information aspect is not peculiar to banks, rather, to any lending relationship. In this sense the information-theoretic approach is concerned with a particular reason of existence of banks and is much narrower in it than the portfolio theories. Summing up, even though the information-theoretic approach enriches the portfolio theory by a discussion of concrete banking activities, including those of screening, monitoring, and liquidity provision, the former offers a limited take both on asset transformation and on *raison d'être* of banking, compared to what can be argued based on the portfolio theory.

There are several features of the portfolio management and information-theoretic approaches that need to be emphasized in light of the above discussion on the transformation of banking.

¹⁴ A similar criticism of the information-theoretic approach, although in a context of theories of financial intermediation based on risk management, not early portfolio theories, was suggested by Scholtens and Wensveen, who argue that even if markets and information were perfect, there would have still existed a need for qualitative asset transformation (Scholtens, Wensveen 2003, p. 23). Thus, a focus on information asymmetry can only explain the brokerage function of financial intermediaries, i.e. their function with respect to uniform, tradable financial services, but not their broader qualitative asset transformation (Scholtens, Wensveen 2003, p. 33). The conclusion is effectively the same as ours: information asymmetries cannot explain existence of financial intermediaries.

A characteristic feature of the early theories treating financial intermediation as portfolio management is that they are broad enough to describe any banking activity. Banks are specialists in asset transformation along assets' liquidity, maturity, and credit risk dimensions. This transformation is performed in a variety of ways, so that any banking activity qualifies for this description. For example, even provision of derivatives can be treated as a form of asset transformation (Boyd, Gertler 1994). Such a theory is incapable of capturing qualitative changes in what banks do, and all the changes are inevitably reduced to a mere change in form how an asset transformation is performed. This theory does not show whether various forms of asset transformation can be qualitatively distinct and have different grounds and sources of remuneration. For this reason even if the portfolio theories continued to dominate the mainstream theories of financial intermediation, it is not clear how they could offer insights relevant for understanding the nature of the transformation of banking. It will become apparent in the discussion of some of the new theories of financial intermediation, in particular those by Gorton (2008, 2009) and Pozsar, Adrian, Ashcraft and Boesky (2010), that can be viewed as a resurrection of the asset transformation approach to banking.

By the same token, the information-theoretic approach to banking treats information asymmetry in a variety of forms as the root cause of all banking-related phenomena and all banking functions. It is information asymmetry that is responsible for screening, monitoring, liquidity insurance, and even for agents' limited participation in the capital markets making financial intermediaries important even in a more market-based financial system. To the extent that information remains asymmetrical, banking does not change. Given that the changes analyzed by the empirical current of the

mainstream literature are not associated with elimination of information asymmetry, banking is still banking. The theory is incapable of capturing other aspects. It is not suited to analyze not only changes in banking within capitalism – those associated with changes not related to information asymmetry, but also its ahistoricism would prevent such a theory from recognizing broader changes, e.g., a difference between capitalist banking and banking in general. At the same time, the information theoretic approach acknowledges that, with a rise in markets as liquidity providers the role of banks in handling information asymmetry on the liability side declines. With the development of information technologies the need for banks' screening and monitoring function, i.e. handling asymmetry on the asset side, is expected to decline, too. Thus, paradoxically, this theory of banking is inherently a theory of a decline of banking, and the theory of financial intermediation is inherently a theory of disintermediation.

Putting these two features of the information-theoretic approach together, it becomes clear that according to these theories, in the course of the empirically observable transformation of banking banks have become “less of the same”. Banking is still banking – qualitatively the same type of economic functions, although on a smaller scale due to a partial substitution by markets. The information-theoretic approach is capable of capturing only quantitative changes in banking, and possible qualitative changes are reduced to mere quantitative ones.

And indeed, a look at a few attempts to address the question of the transformation of banking within the information-theoretic approach proves this conclusion. Securitization is considered under a rubric of a change in costs due to differences in monitoring costs with and without securitization (Mester 1992). A broader approach

towards the transformation of banking treats changes on the liability side and a decline in the role of core deposits as a driving force behind a decline in traditional relationship lending (Berlin, Mester 1998). It shows the interconnection between a decline in the two key dimensions of banking activities. Nevertheless, both aspects are still treated as a quantitative decline in traditional banking functions on the asset and liability sides.

In general, the transformation of banking is addressed within the information-theoretic approach in a one-sided way. Changes in the structure of lending are not considered in the literature at all. And the changes that are addressed show a decline in the traditional banking functions, as in the case of securitization and a decline in deposits triggering broader changes. Specificity of banks' liability side is reduced to a particular instrument, demand deposits proper, and not a social relationship of liquidity provision on demand that can take multiple forms. In a sense, it is a return to a pre-Gurley/Shaw/Tobin state of the debate.

Such a limited treatment of the transformation of banking, with a focus on quantitative changes, is itself problematic. It obviously lags behind the complexities discussed in the mainstream empirical literature on the transformation of banking. Nevertheless, it would not have posed a major problem if the transformation of banking discussed above implying a decline in traditional banking within the information-theoretic framework did not co-exist with a constant share of banks profit in total profits of financial intermediaries. This disjuncture becomes even more severe once one takes into account that banks perform their functions for the economy as a whole, so a more meaningful comparison would be that with total domestic profits. If the information-theoretic approach predicts a quantitative decline in banking and in the role of its old

functions, how can banks receive a rising share of domestic profits? What are the grounds for their remuneration? A too narrow view on the reasons behind existence of banking characteristic of the information-theoretic approach makes itself visible in the inability of this theory to address these questions. Is this rise in profits associated with new functions not captured by the old theories? Or with new ways of performing old functions, not recognized by the information-theoretic approach?

To the best of our knowledge, the paradox of a rising bank profit share co-existing with a decline in traditional banking functions as they are understood within the mainstream theory has not been explicitly stated in the literature. Nevertheless, a similar paradox drew attention of Allen and Santomero (1997). They noticed that competition for brokerage fees since the early 1970s resulting in a decline in trading costs for individuals did not lead to a decline of mutual funds, and that technological revolution reducing costs of information and information asymmetries did not lead to a shift to direct market participation. Thus, a decline in market imperfections that are usually viewed as *raison d'être* of financial intermediaries did not make them disappear.¹⁵ This obvious evidence of the information-theoretic approach looking for grounds for financial intermediation in a wrong place made Allen and Santomero (1997, 2001) call for a new theory of financial intermediation.

¹⁵ This argument by Allen and Santomero resembles Nell's criticism of a somewhat broader proposition by Hicks. Hicks suggests that a possible way of looking at monetary history is through the lens of institutions emerging to lower transaction costs (Hicks 1967, p. 7). Nell argues such an approach is misleading, in light of the rising share of the financial sector in GDP in the second half of the 20th century (Nell, 2001, p. 188-189).

2.6. Mainstream Theories Reconsidered: Theories of Transformation of Banking, or A Quest for a New Content of Banking

The inherent limitations of the existing theories making them incapable to address some of the issues of the transformation of banking and reconcile an apparent decline in traditional banking with rising importance of financial intermediaries gave rise to a quest for new theoretical approaches to banking. This search was reinforced by an outbreak of the crisis that started in 2007 and had some features of a banking crisis, but at the same time had its specificity making it different from bank runs as we know them from history. Emerged to explain the change, the new theories are by definition better suited for an analysis of the transformation of banking.

A reconciliation of co-existence of the transformation of banking with its remaining importance required an argument about emergence of new functions of banking or new ways of performing old functions. Otherwise, one could not explain rising profits accruing to financial intermediaries and rising value added of them. Thus, there emerged four distinct approaches belonging to new theories of financial intermediation. Each of these approaches has its own take on what constitutes the essence of banking and its implications for the transformation of banking. Consider each of these approaches in detail.

Probably the most widely known new theory of financial intermediation is the one suggested by Allen and Santomero (1997, 2001). They argue that the information-theoretic approach focuses on functions of banking that are not relevant any more. Banks have transformed, and they perform two main functions – risk management and risk trading, and facilitating participation of others by reducing participation costs through

creating products with stable distribution of cash flows. These two functions have become the core of banking due to a rise in the market-based financial system and due to existence of higher participation costs for entities other than financial intermediaries. Not only is risk management the major activity of intermediaries, but also financial intermediaries are the main entity in the economy performing this function. They are uniquely suited to perform this function, because they are specialists dealing in financial assets, and any dealing in financial assets – origination, trade, servicing – by definition involves managing and trading financial risk (Allen, Santomero 1997, p. 1478). As a result of this transformation of financial intermediaries, a traditional distinction between those and the markets has broken down, as the markets are now dominated by intermediaries¹⁶.

Incidentally, this approach pivoting on risk was anticipated by Alan Greenspan as early as in 1994. He defined “traditional banking” as “the measurement, management, and acceptance of risk” (Greenspan 1994, p. 3), by linking traditional banking functions to risk. For him, liquidity insurance is based on risk diversification through pooling assets, and payment services involve accepting and managing risk (Greenspan 1994, p. 3). Contrary to Allen and Santomero (1997, 2001) who stress risk only in light of the transformation of banking, for Greenspan, the basic functions of banks understood in this way remain unchanged, although technological characteristics of products, specific means by which these functions are performed, and precise character of the risks

¹⁶ From a theoretical viewpoint it is an interesting statement, as it implies that capital markets are akin to the money market that is created by and exists for intermediaries, not firms or individuals.

involved have evolved¹⁷. On this basis he concludes that the conventional wisdom that banking is a declining industry “may reflect too narrow a vision of the role of banks” (Greenspan 1994, p. 2).

Allen and Santomero’s approach was critically appraised by Scholtens and Wensveen (2000, 2003) who argue that Allen and Santomero are correct in their emphasis on risk, but do not go far enough with their argument. For Scholtens and Wensveen, risk management is what defines banking, its main *raison d’être*, it is not a new function of banking due to its transformation. The origin of banking is in its risk transformation and risk management activities, since banking in the Italian Renaissance¹⁸, – a somewhat similar argument to Greenspan’s. Reducing costs and informational asymmetries are a part of the banking process, “but it occurs as a by-effect” (Scholtens 2000, p. 1251): financial intermediaries “deal in money and in risk, not in information *per se*. Information production predominantly is a means to the end of risk management. In the real world, borrowers, lenders, savers, investors and financial supervisors look at them in the same way, i.e. risk managers instead of information producers” (Scholtens, Wensveen 2003, p. 23, see also p. 21, 27)¹⁹. Finally, the other extension of Allen and

¹⁷ It is interesting to note that, for Greenspan, within the range of functions related to risk, “risk information processing” has become increasingly important and “now lies more visibly closer to the core of the banking business” (Greenspan 1994, p. 4). Thus, although Greenspan’s emphasis on risk in many ways anticipates Allen and Santomero’s argument, on the other hand, it is a reversal of their approach, with a shift from handling risk as a generic function of banks to information-related activities linked to handling risk. This seems to be an attempt by Greenspan to combine his reasoning with the information-theoretic approach dominant at that time, whereas Allen and Santomero stress precisely the opposite transformation – away from information-related services to risk management.

¹⁸ A problem with Scholtens and Wensveen’s treatment of risk is that they use categories of risk transformation, absorption, management, and trading, with these four conceptually distinct categories being mixed together, so it remains unclear what are the exact functions of banks with respect to risk.

¹⁹ Scholtens and Wensveen develop this argument into a broader – and quite powerful – criticism of the information-theoretic approach, arguing that it is good at analyzing specific aspects and a variety of contingencies of financial decisions, but its power is also its weakness. The theory becomes a set of *ad hoc*

Santomero's approach is in a model of delegated risk management developed by Hakenes (2004).

The second approach within the new theories of financial intermediation is represented by Rajan (1996) arguing that the essence of banking is *liquidity provision* on both sides of the balance sheet. For him, it is not incidental that banks historically fund themselves with demand deposits and extend loans, as both activities are effectively liquidity provision on demand (Rajan 1996, p. 119). Their co-existence can be explained by economy of scale and co-insurance – a situation when holding illiquid assets is required to maintain credibility of liquid liabilities, due to the paradox of liquidity discussed above. In light of this theory, the transformation of banking is merely a change in “the outward form of banks’ activities – though not their underlying economic function” (Rajan 1996, p. 114). The transformation consists in banks’ concentrating on the essentials of their function, ridding of the rest. An example is issuing letters of credit instead of a direct funding of a loan, reflecting the essential bank function as provision of “not funding per se, but funding *on demand*” (Rajan 1996, p. 121-122, italics in the original).

The other two new theories of financial intermediation emerged largely as a response to the 2007-2009 crisis. The motivation behind them is to define the essence of banking in a way that would allow for an analysis of shadow banking as real banking.

The first of these approaches is proposed by Gorton for whom banking is creation of informationally-insensitive debt – debt, value of which is not sensitive to private

findings having anecdotal value, with no general and coherent explanation of financial intermediation (Scholtens, Wensveen 2003, p. 21-22).

information (Gorton 2009, based on Gorton, Pennacchi 1990)²⁰. The relevant characteristics of this type of debt are that it has no fixed maturity, hence, is redeemable at par on demand; it can be used in transactions (serves as a means of circulation, transferrable to other parties); it is senior and backed by a portfolio of diversified assets (Gorton 2009, p. 14). The first historical form of informationally-insensitive debt is demand deposits, but repo possesses all the four characteristics of the informationally-insensitive debt, as well, hence, can be viewed as a wholesale form of deposits redeemable on call. On these grounds, Gorton concludes that shadow banking is real banking. Furthermore, shadow banking as “creating a “currency” for firms” (Gorton 2009, p. 41) is grounded in material conditions, for securitization is a necessary development due to a rise of demand for collateral and for wholesale deposits of large firms, both financial and non-financial (Gorton 2009, p. 39-40)²¹. Incidentally, Gorton’s treatment of the 2007-2009 crisis as a run on repo, i.e. a panic on the wholesale market, was anticipated by Boyd and Gertler already in 1993, who argued that the rising exposure to liquidity risk in the course of transformation of banking can lead to a banking panic on the money market (Boyd, Gertler 1993, p. 328). For them, a first sign of bank panics hitting the wholesale market was at the time of the collapse of Continental Illinois in 1984 that triggered a panic withdrawal by large CD holders.

²⁰ There are some deeply rooted problems with this approach, as it requires exogenous shocks to explain how informationally-insensitive debt becomes informationally-sensitive in the time of a bank panic. It effectively implies – without Gorton’s acknowledging it – that information sensitivity is not a feature of debt per se, but of social relations, trust, etc, i.e. that it is not a private, but a social phenomenon. But Gorton’s approach cannot reconcile this feature of bank debt as a social phenomenon with an analysis of banks as private institutions.

²¹ As he put it, “Forcing everything back on balance sheet seems like an attempt to return to the Eden of the Quiet Period without recognizing that the world has changed... [T]he rise of securitization and repo as very significant parts of the capital markets cannot be explained as a bubble, or as the product of greed, and so on” (Gorton 2009, p. 40).

Finally, the other attempt to reconceptualize financial intermediation is made by Pozsar, Adrian, Ashcraft, and Boesky (2010), who define the essence of banking as credit intermediation through credit, maturity, and liquidity transformation. This approach effectively represents a rebirth of Gurley and Shaw's asset transformation as the essence of banking, without its origin being tied to information asymmetry and market imperfections. Pozsar and his colleagues argue that shadow banking plays an analogous economic role as traditional banks – that of credit intermediation (Pozsar et al. 2010, p. 13) – in which credit, maturity, and liquidity transformations are unbundled and performed through steps in a chain of specialized institutions, when the traditional credit intermediation process is “vertically sliced” (Pozsar et al. 2010, p. 11, 69, 72). It is the shadow credit intermediation process, which has lengthening of the intermediation chain as its counterpart²².

There are two important theoretical implications of this approach.

First, the separation of the liquidity, maturity, and credit risk transformations across steps and institutions within the shadow banking system reveals that these are inherently different processes, “three independent concepts” that do not need to be performed simultaneously, although it used to be the case in traditional banking (Pozsar et al. 2010, p. 18).

Second, it makes the authors conclude that the nature of lending has been changing (Pozsar et al. 2010, p. 28-29). Lending is no longer reliant on banks only, rather on a network of institutions funded through wholesale market and capital markets

²² Lengthening of the intermediation chain as one of the key features of the transformation of banking was also noticed by Samolyk (2004).

globally²³. Although banks' direct involvement is only at the level of loan origination, their indirect participation is much broader. Bank subsidiaries of bank holding companies are effectively lender of last resort to their non-bank subsidiaries²⁴. Finally, lending is capital efficient but depends on liquid wholesale funding and global capital markets, hence, can easily become capital deficient.

In their discussion of the origin of shadow banking, Pozsar and his colleagues treat shadow banking as a unity of "parallel banking" and shadow banking proper (Pozsar et al. 2010, p. 45, 67-68, 72). The former has arisen due to gains from specialization and comparative advantage over traditional banks, with most of the institutions of the parallel banking to be found in the "external" shadow banking segment – diversified broker-dealers (DBD) and independent specialists. The latter has originated in regulatory arbitrage and has limited economic value²⁵.

There are two main strengths of the new theories of financial intermediation. First, by showing an emergence of new banking functions (as in the case with Allen and Santomero's treatment of risk) or by arguing that old functions (and, hence, old grounds for remuneration) are preserved and merely take new forms (understood as liquidity provision, asset transformation, or creation of informationally-insensitive debt), the new theories of financial intermediation offer an answer to the question how the

²³ Although lending hinging on a broader set of institutions than banks proper becomes obvious with shadow banking, it is inherent in the very nature of lending based on reallocation of idle funds. In this sense securitization is merely a reaffirmation of the social character of loanable capital, not a qualitative change in the nature of lending, as Pozsar and his colleagues try to argue (Pozsar et al. 2010, p. 28).

²⁴ It is similar Samolyk's argument that layering of intermediation makes it harder to quantify the role of banks (Samolyk 2004, p. 31, 35, 51).

²⁵ It looks like their argument is that these two forms are different in their origin, but both are forms of credit intermediation, regardless of what caused its genesis.

transformation of banking can co-exist with a rising bank profit share. Thus, the new theories resolve the main problem which could not be reconciled by the second wave theory of financial intermediation.

Secondly, and broader, the new theories of banking break with the limits of the information-theoretic approach discussed in the previous part of this chapter, where it was compared to the earlier portfolio theories. Information asymmetry and market imperfections do not seem to be at the core of existence of financial intermediaries that rather exist to manage risk, transform assets, or create a specific type of debt – informationally-insensitive. None of these aspects has an immediate connection to market imperfections, none of them uses a perfect market as a reference point, hence, the new theories represent a broader approach to what banking is.

In spite of these obvious advantages of the new theoretical conceptualizations of banking over the information-theoretic approach, they also have two major limitations.

First, several dimensions of the transformation of banking that were not addressed by the information-theoretic approach are not discussed by the new theories either. Among these aspects are implications of lending to different economic sectors and relationship between bank revenues and profits and those of the rest of the economy. It reveals that in spite of the major differences among the “old” and the “new” mainstream theoretical approaches, they share a common feature of not distinguishing across economic sectors outside of banking. Thus, all of them pivot on what can be called “banking view” of financial intermediation, not a macroeconomic or a political economy view of the society as a whole.

Second, an important aspect of all the theoretical approaches to the transformation of banking, both within the information-theoretic framework and new theories of banking, is that all of them either deny the transformation as a qualitative change in banking, arguing that its functions have remained the same and merely took new forms, or show that the content of banking has undergone a major change. Examples of the former are Rajan's emphasis on liquidity provision, Greenspan's, Scholtens and Wensveen's treatment of risk. The second approach is best represented by Allen and Santomero's take on risk management. Thus, the range of approaches to the transformation of banking manifests prevalence of a "dichotomous thinking": banking is either totally different today, or exactly the same, with only a concrete way they perform their functions having changed. What at first glance appears to be an exception, for instance, approaches developed by Gorton and Pozsar and his colleagues, at a closer examination turns out to belong to the first group of theories, stressing new modes of performing the same set of functions. In the case of Gorton and Pozsar, it is, respectively, creation of informationally-insensitive debt and credit intermediation through liquidity, maturity, and credit risk transformation, merely changing their concrete forms.

From the perspective of this study, contrary to a dichotomous thinking of the dominant theories pivoting on the "change vs. no change" debate, the question of the transformation of banking is more complex. Instead of such dichotomous thinking, there is a need for a theory of banking that would allow for both a trans-historical content of banking and changes in banking, with these changes being more than just a change in a concrete form of performing the immutable functions. Chapter 4 suggests a possible theoretical framework for theorizing banking. A Marxist theory of banking will be shown

to have a dual foundation – a core function of banking seen as liquidity provision through exchange of promises to pay and an explicit connection between banks and flows and stocks of value. Such dual foundation allows for a relatively unchanged core function of banking to co-exist with and even form a basis for a significant transformation of banks’ relationship to the rest of the economy and, consequently, for a non-trivial change in the character of bank revenues. It is therefore the very nature of what banking is – which remains the same – that allows for its changing content through history. Thus, a change in banking is intrinsic in what banking is.

2.7. Conclusion

The mainstream empirical literature acknowledges the transformation of banking since the 1980s. Nevertheless, its treatment of this transformation is incomplete, as it does not consider differences in lending to different economic sectors. It reveals that the approach is based on a “banking perspective” – for a bank it does not matter who a loan is extended to. But it does matter from the perspective of the society as a whole, hence, from a broader political economy perspective.

In spite of this major limitation, the empirical discussion of the transformation of banking points out two relevant facts that are also important to the present study. First, it shows that profit is a more adequate measure of banks’ market share, not their assets. It can be used as an empirical basis for a conclusion that banks are not dying out. On these grounds, the present study will focus on banks’ profits and revenues as a measure of banks’ role. Nevertheless, given that banks perform their functions with respect to the economy as a whole, we will develop this argument further and consider banks profit

share in total domestic profits, not profits of financial intermediaries, as in the mainstream literature. Secondly, one of the empirical studies correctly points out that not only is profit an adequate measure of banks market share, but also it is a more appropriate lens for an analysis of the ongoing transformation of banking, due to the nature of this transformation itself. On this basis, we will focus on banks income statements to analyze the content of the transformation of banking.

The empirically acknowledged transformation of banking poses a challenge to the dominant mainstream theory of banking. The information-theoretic approach considers this transformation as a merely quantitative change, with banks becoming “less of the same”. Not only is it not capable of capturing a qualitative shift, but also the major conclusion of a quantitative decline cannot be reconciled with a rising bank profit share in total domestic profits. This, coupled with a limited treatment of the transformation itself, reveals weaknesses of the existing theory, which turns out to be incapable of explaining the major structural changes in banking. For this reason, to be able to explain changes, the theory of financial intermediation itself needs to be changed.

New mainstream theories of banking emerged to fill the gap by redefining the content of banking along several lines. Among them are liquidity provision on demand on both sides of the balance sheets, informationally-insensitive debt creation, asset transformation, and risk management and trading. The disjuncture between the theoretical and empirical strands of the mainstream literature has been narrowing with an emergence of these new theoretical takes on banking.

Although the new theories are more adequate to explain the change and pose banking on a broader foundation than mere market imperfections, changes in the

structure of lending and broader implications of changes in the mode of banks funding themselves cannot be addressed by either old or new theories, because of their internal limitations. The reason behind it is that the mainstream theories of banking fail to locate it within a broader macroeconomic theory or a theory of capital as a whole.

From a political economy perspective, a theory of banking capable of capturing these issues needs to have an explicit account of the relationship between banking and the circuit of capital and revenue with underlying social relations, and an explicit treatment of the nature and origin of profits in its relationship with flows of value. That is, banking needs to be put into the context of the capitalist mode of production and specifically capitalist sources of profit in value creation. It is a distinct feature of a Marxist theory of finance that locates the credit system within the capitalist mode of production, both in its genesis and functioning. Such a theoretical framework would have several implications.

An immediate consequence of this approach to banking would be its ability to address the aspects of the transformation that cannot be captured by the mainstream theories, but that are relevant from a macroeconomic and political economy perspectives. As a result, the content of the transformation of banking would be sought not in a changing mode of performing the same set of core functions, but in a changing place of banking with respect to value creation and distribution, changing relationship with the circuit. Thus, it is a relationship of banking to the circuit of capital and revenue that gives meaning to banking and its transformation. This is missing in the mainstream theories.

A corollary to this is that the transformation of banking understood in this way could be shown to be connected to changing sources of profit as related to flows of value – past, present, and future. Hence, we will look into the link between the transformation

of banking and the transformation of the content of bank profit. Where do profits from the transformed banking come from? This question becomes even more relevant in light of the transformation of banking co-existing with a rising bank profit share in total domestic profits.

A broader implication of this approach to banking is that it would allow for a co-existence of a trans-historical content of banking and changes in banking, going beyond the dichotomous thinking of the mainstream theories.

CHAPTER 3

BANKING, TRANSFORMATION OF BANKING, AND BANK PROFIT IN THE THEORY OF THE MONETARY CIRCUIT

3.1. Introduction

It was argued in the previous chapter that banks have been changing over the past few decades. This transformation has posed a challenge to the mainstream theories of banking that proved incapable of fully capturing this transformation. The information-theoretic approach has encountered its limits, and new theories of banking have emerged posing banking on a foundation other than information asymmetries and market imperfections. In spite of being better suited to address the transformation of banking, the new theories fail to capture some relevant aspects of the transformation of banking. Specifically, both the old and the new theories do not recognize the differences across the economic sectors, hence, fail to posit banks in their relationship with these sectors and account for the related transformation.

Therefore, a need for an alternative theory of banking remains. Such a theory should pose banking on a foundation other than information asymmetries, like the new mainstream, approaches, yet explicitly locate banks among the major economic sectors. The monetary circuit of production theory, also known as the circuitist approach, at a first glance seems to meet these requirements. The present chapter will focus on this theory with an aim to establish to what extent it can provide a broad enough theory of banking to address the ongoing transformation. Special attention will be given to, first, the strengths of this theory compared to the mainstream approaches and the insights it offers that need to be retained. Second, in the course of the discussion limitations of the monetary theory

of production approach will be uncovered. It will be argued that these limitations prevent the theory from offering a foundation for an alternative theory of banking.

The chapter is organized as follows. Section 2 locates the circuitist theory in the history of economic thought. It discusses the main cornerstones of the theory of banking in Schumpeter and Keynes – the immediate predecessors of the theory of the monetary circuit and, more broadly, of the heterodox theories of banking in general. Section 3 focuses on the key ideas of the circuitist theory of banking and shows how the theory addresses the question of the transformation of banking. It is argued that in spite of posing banks on a foundation other than information asymmetry and acknowledging differences across economic sectors, therefore, being well suited for studying the transformation of banking, the circuitist theory nevertheless fails to adequately capture this transformation. It happens due to internal limitations of the theory. Section 4 focuses on the treatment of bank profits within the monetary theory of production. The circuitist take on profits is shown to be problematic in general and even more so in the context of the transformation of banking. Section 5 shows that some insights into bank profits can be gained from a broader discussion of financial profits in the heterodox literature on financialization. Section 6 concludes by formulating requirements to an alternative theory of banking that would retain the insights of the circuitist approach to banking and the discussion of financial profits in the debate on financialization, but at the same time overcome the limitations of these approaches.

3.2. Schumpeter and Keynes: Immediate Origins of the Circuitist Theory of Banking

The previous chapter has concluded that a theory of banking capable of capturing its ongoing transformation should pose banks on a foundation other than information asymmetry and market imperfections, similar to the new mainstream takes on banks. At the same time, contrary to both the old and the new theories of banking, it should acknowledge differences across economic sectors and explicitly locate banks among those sectors. At a first glance the monetary theory of production developed by Graziani (2003), Parguez (2004, 2001, Parguez, Seccareccia 2000), and Bossone (2001a, 2001b, 2003) seems to possess both characteristics, hence, appears as a good candidate for such a theory. Before discussing this theory and the way it treats the transformation of banking in a greater detail, consider its origin. Approaches to banking developed by Schumpeter and Keynes can be viewed as immediate predecessors of the heterodox theories of banking, in general, and the circuitist theory, in particular, therefore, they deserve a closer analysis.²⁶

3.2.1. Schumpeter: Banking for Development and Creation of Purchasing Power

Schumpeter can be viewed as one of the immediate predecessors of the modern heterodox theories of banking and especially the circuitist theory. His approach is based on two main pillars – an emphasis on banking as a medium for economic development and on banks as essentially specialists in money creation (Schumpeter 1961

²⁶ Although one might question whether Schumpeter and Keynes should be considered part of heterodox economics in general, when it comes to theories of banking, their analysis is distinct enough to make them stand apart and be a part of the heterodox tradition.

[1911/1934]). Both aspects have played an important role in the modern theories of banking, nevertheless, Schumpeter's originality should not be overstated. An emphasis on development banking can be traced back to Sir James Steuart (1770b, book IV, part II), and the focus on money creation by banks can be found in all the monetary theories existing before Schumpeter, although he was among few who would define banks as primarily money creators. Elsewhere Schumpeter (1954, p. 317-318) identifies the origin of this view in the work by Daniel Webster who made note issue the defining trait of a bank as early as in 1839. Later, this approach was also held by Sidgwick, Fetter, Withers, and Hahn, among others.

Schumpeter studies credit and banking in their relationship with development. The primary focus of his analysis is "spontaneous and discontinuous changes in the channel of the circular flow" (Schumpeter 1961 [1911/1934], p. 65), not so much because this is the way economic changes have always occurred, but more because of their "fruitfulness" (Schumpeter 1961 [1911/1934], p. 63). He argues that carrying out new combinations of existing resources, i.e. the process of development, requires purchasing power. Thus, for Schumpeter, credit is defined through the requirements of development. He sees credit as essential for one's becoming an entrepreneur. At the same time, it is not fundamental for all the other types of economic activities – consumption, maintenance of a disturbed business, and running business in the already established circular flow (Schumpeter 1961 [1911/1934], p. 102-103). These other activities can be theorized prior to and independent of credit, whereas development cannot, therefore, Schumpeter focuses on credit for development and concludes that the two are inseparable.

Further, credit for development cannot come from a reallocation of the already existing purchasing power, because it is committed to the already existing circular flows (Schumpeter 1961 [1911/1934], p. 72, 96). Therefore, credit in Schumpeter is based on creation of new purchasing power. To start a new flow, new liquidity is required. This in turn serves a basis for Schumpeter's theory of banking. For him, "creation of means of payment centres in the banks and constitutes their fundamental function" (Schumpeter 1961 [1911/1934], p. 98, see also p. 185). A bank is a producer of commodity "purchasing power" (Schumpeter 1961 [1911/1934], p. 74). In Schumpeter's view, the emphasis on creation of purchasing power is fundamental both at a theoretical level for the reasons just discussed, and at an empirical level, because according to his estimates in his time in some countries three-quarters of deposits were credits created by banks themselves (Schumpeter 1961 [1911/1934], p. 99). And even if they are not, following Hahn²⁷, Schumpeter treats the actual deposits as a particular type of bank credit – the one without a corresponding increase in depositor's purchasing power.²⁸

An implication of such a take on banking is that for Schumpeter the functions of banks not related to creating new purchasing power are secondary, derivative (Schumpeter 1961 [1911/1934], p. 125, 189, 199-200). For instance, banks also advance credit for purposes other than development, on the one hand, and assemble the already existing liquidity scattered across the society, on the other hand. But for Schumpeter

²⁷ According to Robertson (1937, p. 430), a similar idea was expressed by Harrod in *The Trade Cycle*.

²⁸ From a bank's perspective, lending out funds acquired through collecting deposits and lending out a claim on itself are obviously equivalent. The same holds for a borrower. Nevertheless, from a political economy perspective, this cannot be a basis for denying deposit collection as a separate function of banks, different from issuing claims on itself, because the former puts banks in a position of centralizing idle funds scattered across the society, whereas the latter does not.

these functions of banking are secondary with respect to development banking, because it is development that gives meaning to transactions related to the circular flow. On the one hand, credit for maintaining the established circular flows becomes important only after these flows already exist, which happens as a result of development. With no development, there would be no circular flows. On the other hand, only an established circular flow can release temporarily idle funds. On these grounds Schumpeter concludes that forms of credit not related to development and credit based on intermediating existing, instead of the newly created, liquidity are a consequence of development and become the secondary function of banks. Banking as creation of new purchasing power should be theorized before banks as financial intermediaries. “The banker, therefore, is not so much primarily a middleman in the commodity “purchasing power” as a *producer* of this commodity... He is essentially a phenomenon of development” (Schumpeter 1961 [1911/1934], p. 74, italics in the original).

This theory of banking leads Schumpeter to a particular theory of bank profit. There are two aspects of this theory. According to one, interest accruing to banks is kind of a tax on entrepreneurial profit (Schumpeter 1961 [1911/1934], p. 175, 210). The other aspect of interest is its being a premium on present over future purchasing power, or a premium on newly created purchasing power (Schumpeter 1961 [1911/1934], p. 157, 187-188, 196). In both cases interest comes from the surplus value created in the course of development, i.e. from profit from carrying out new combinations (Schumpeter 1961 [1911/1934], p. 173-175). Therefore, it involves a price struggle between a borrower and a lender (Schumpeter 1961 [1911/1934], p. 125, 192-193).

Schumpeter's emphasis on the role of banks for economic development is important, as it shows that banks can be a powerful medium in this process. Nevertheless, by linking banking and development, Schumpeter ends up developing a partial theory of banking, which, contrary to his belief, is a poor basis for generalizations to capture other aspects of banking activities. And indeed, how can banking for development serve a basis for understanding the ongoing transformation of banking as described in the previous chapters? How could it nest rising lending to households? How could it explain securitization and other non-lending bank activities? In this sense, Steuart's approach to banking is more powerful, because although he also stressed a special role of development banking, he was clear from the outset that there are different generic types of banks, with each having its own principles of operation that cannot be deduced from any particular type of banking (Steuart 1770b, book IV, part II, chap. III).

From the perspective of the present study, even though banks do have a capacity to create money *ex nihilo*, and even though this process is relevant for development, nevertheless, neither development, nor money *ex nihilo* can be a starting point of an analysis of banking. Consider this argument in a greater detail.

First, although Schumpeter is right that under certain conditions banking might be a necessary condition for development, the converse is not true²⁹. Development is neither

²⁹ Here we are not discussing a broader problem with Schumpeter's approach – to what extent development is a good organizing principle for an analysis of capitalism. It creates an illusion that the fundamental driving force of capitalism is the pursuit of development. Schumpeter validates his entry point emphasizing “fruitfulness” of development (Schumpeter 1961 [1911/1934], p. 63) but it does not resolve the problem. His approach is misleading, because, first, it pivots on an abstract process of economic development, without studying social relations involved. There can be different types of development, but all of them are subsumed by Schumpeter under an abstract process of development. As a result, the industrial revolution in England and economic development in the Soviet Union in the 1920s might appear not that different. Second, without being explicit about it, Schumpeter focuses mainly on the capitalist type of development, but even that alone cannot be studied as an abstract process of changes in the circular flow. It itself should rather be viewed as a consequence of the drive for profit, i.e. for self-expansion of value as a fundamental

a necessary, nor a sufficient condition for banking, therefore although looking at banks in the context of development might offer insights into the specificity of banks providing credit for development, and can explain why and how banks can be a powerful medium for this process, it cannot be a starting point for developing a general theory of banking, contrary to Schumpeter's view.

His vision that a general theory of banking can start with banking for development reveals a deeper methodological problem in Schumpeter – a problematic take on what is abstraction. His abstraction is “mechanical”, formal. He isolates a new circular flow from both other circular flows co-existing with it and from the preceding process of accumulation. But this is not a fruitful approach, as, first, the current circular flows can release temporarily idle purchasing power that can be channeled into development. That is, although Schumpeter is right that the existing purchasing power is already absorbed by the established circuits, it does not prevent these circuits themselves from releasing monetary funds. Second, for historical reasons it is also misleading to assume away the purchasing power accumulated in the past, as in many cases development has been carried out by means of monetary wealth concentrated in the hands of a specific group of people. Not only new combinations in the already developed capitalism, but even the very beginning of the capitalist development is usually based on the previous economic systems, domestic or foreign, so assuming away the results of the past accumulation is misleading both historically and logically. One example would be

capitalist process. It is not development, but value expansion, that constitutes the fundamental capitalist phenomenon. Put differently, new enterprises are established not because capitalists seek to carry out new combinations, but because carrying out new combinations is a necessary condition for profit extraction. And in this sense development is only a means for and a consequence of profit extraction. Thus, starting with development is in many ways analogous to starting an analysis of capitalism with use values, which obscures the driving force of capitalist production – an approach criticised by Marx.

the monetary wealth accumulated by merchants and money-dealers from their previous operations that was channeled into production during the industrial revolution. In this sense an attempt to establish an abstract first circuit that would be a basis for all the other circuits is deeply problematic. Therefore, it is misleading to cut a new circuit off from other existing circuits and from the results of the previous accumulation.

This is not to deny that Schumpeter is right that to study banks systematically one needs an abstraction. Nevertheless, this abstraction should not be formal. A better abstraction would be the simplest form of the banking relations that contains a possibility to be developed into other forms of banking activities.³⁰ As it will be argued in chapter 4, such an abstraction could be a credit relationship in general and the bank business of discounting, not the process of development.

Schumpeter's mode of abstraction is interwoven with other aspects of his theory. Development banking and money ex nihilo are not two independent characteristics of banking – both of them rather stem from the nature of his abstraction. Starting his analysis with development makes Schumpeter not only treat current credit as secondary, but also focus on money creation ex nihilo necessary. Therefore, once the problematic nature of his abstraction is uncovered, it becomes clear that development banking cannot be a starting point for a theory of banking. By the same token, money ex nihilo – the other pivotal concept in Schumpeter's theory of banking – is also a problematic starting point for an analysis of banking, for a number of reasons.

³⁰ This take on abstraction is based on the work of the Soviet philosopher Evald Ilyenkov for whom an abstraction, or what he called “the concretely universal”, is “the genuinely universal foundation of the whole system, its ‘elementary cell’” (Ilyenkov 2008 [1960], p. 80). An abstraction should contain, “like a ‘cell’ or embryo, the wealth of more complex, more developed forms of... relations” (Ilyenkov 2008 [1960], p. 85).

If development were not to be used as the entry point, therefore, if there were room for other sources of funds, demand for liquidity could be equally met by intermediating existing liquidity or creating new liquidity. It was Keynes' argument who, acknowledging the role of initial finance in starting a circuit, insisted that demand for finance can be satisfied by either creating new liquidity or intermediating existing liquidity (Keynes 1937a, p. 664-665, 1937b, p. 245-246, 1938, p. 319). Trying to preserve Schumpeter's emphasis, Graziani explains Keynes' focus on both forms of liquidity provision by his attempt "to modify as little as possible the approach followed in the *General Theory*, where the stock market occupies the main position, and the banking system seems to be entirely absent" (Graziani 1987, p. 36). From our perspective, Graziani's explanation is misleading, because a casual observation alone shows that banks extend credit based on both deposits they collect and by creating bank money. And the fact that all modern money is credit money does not alter the principles of functioning of banks.

Furthermore, not only can liquidity indeed be provided through these two channels, as argued by Keynes, but also the very question of whether banks provide liquidity based on deposits or issuing notes can be posed only after a discussion of general principles of bank credit. Put differently, the questions of bank credit and its concrete mechanisms emerge at different levels of abstraction. There are two approaches to this. According to a Marxist theory of banking as developed by Itoh and Lapavistas (1999, p. 94-95) and Lapavistas (2003, p. 79), banks develop functions of collecting deposits and issuing notes as means of increasing their lending capacity, therefore, of raising profits. In this theory, banks are both intermediaries and money creators, and there

is no hierarchy between these two functions, although both of them are viewed as concrete forms of doing banking business, not the content of this business in general.

Another approach was suggested by Hicks. For him banks in general are dealers in securities, both marketable and unmarketable (Hicks 1967, p. 48), or dealers in money (Hicks 1989, p. 67). Similar to Itoh and Lapavistas, deposit taking and issuing notes are more concrete forms of doing the banking business. Nevertheless, Hicks also acknowledges hierarchy between these two functions, and for him issuing bank notes develops out of deposit taking. He shows that a theorization of banks as liquidity creators requires a prior theorization of demand deposits withdrawable at sight (Hicks 1989, p. 55-58).³¹ For Hicks, banking spontaneously arises from discounting promissory notes which in turn serves a foundation for attracting deposits. A need to expand banking business of discounting and lending in general makes banks actively encourage deposits through paying interest on demand deposits and making them transferrable, i.e. facilitating payments. As a result of deposits being withdrawable at sight, both by depositors and their creditors, demand deposits become money. It is this moneyness that becomes a foundation on which lending out claims on bank's own debt becomes possible. Hicks stresses that the act of bank lending by creating a deposit appears differently to different participants of this transaction (Hicks 1989, p. 58). For a borrower, this type of borrowing is not any different from borrowing cash. This is why moneyness of bank deposits is crucial. For a bank, this corresponds to a simultaneous increase in its assets and liabilities. And only from the perspective of the society as a whole, bank creates

³¹ Incidentally, a similar argument was implicitly suggested by Keynes when he talks about bookkeeping nature of liquidity provision (Keynes 1937a, p. 666, 1937b, p. 247). It was later also advocated by Chick (1992, p. 194-195) in her theory of evolution of banking, according to which at an early stage banks were pure intermediaries and later evolved into entities issuing debts on themselves.

money. Thus, for Hicks, lending by creating deposits is not only a particular form of lending, but also a derivative form that can be theorized only after demand deposits have become money. Put differently, creation of purchasing power ex nihilo cannot be viewed as the simplest abstraction containing more developed forms of lending in an embryo form. Banks' deposits being redeemable on demand forms a foundation for banks' ability to lend out own debt. Thus, banks do not emerge as entities creating money, hence, cannot be defined on these grounds. Rather, as Nell put it in his account of the evolution of Hicks' monetary thought, banks "deposits can be used as means of payment; then banks come to create money" (Nell 2001, p. 187).

This has important implications for Schumpeter's theory of banking. Regardless of a hierarchy between deposit taking and money creation, the approaches discussed above suggest that these two functions accrue to banks at a later stage of their development and constitute concrete ways of doing the banking business. Therefore, an attempt to develop a theory of banking based on money creation could not possibly lead to a general theory of banking. Schumpeter's emphasis on money creation would not have been a problem, if it did not involve relevant implications for heterodox theories of banking. But it does, as the emphasis on money creation ex nihilo does not allow one to pose, let alone resolve, a fundamental question of the social relations behind acceptability of bank liabilities as money, hence, of the social relations behind banking in general. Money creation becomes an abstract process without any content.

To recap, Schumpeter shows that banks are relevant for development and discusses one of the possible ways they can provide liquidity, but it does not answer a question what a bank is and why banks exist. Hence, his theory of banking would

necessarily be one-sided, therefore, it is not surprising that it does not seem to be of direct value for understanding the ongoing transformation of banking and bank profits since the 1980s.

3.2.2. Keynes: Banks and Liquidity Provision

Keynes can be viewed as another immediate predecessor of the modern heterodox theories of banking. He has not developed a theory of banking per se, nevertheless, his writings on monetary issues offer some insights relevant for understanding the nature of the banking business.

For Keynes, banks are essentially liquidity providers.³² They specialize in organizing and managing “a revolving fund of liquid finance” (Keynes 1937a, p. 666). For the purposes of the present study it is important that banking is a monetary phenomenon, which implies that information and risk cannot be a starting point for studying banks. Therefore, Keynes’ emphasis on liquidity provides a foundation for an analysis of banking other than information asymmetries and market imperfections.

The emphasis on liquidity is one of the main contributions by Keynes to the history of economic thought. For him, liquidity is paramount for a smooth functioning of an economy, with its absence being able to trigger an economic slowdown or even a crisis. He summarized this insight in his famous conclusion that “the investment market can become congested through shortage of cash. It can never become congested through shortage of saving” (Keynes 1937a, p.669). Given that banks specialize in liquidity provision, hence, can to some extent control it, it places them in a strategically important

³² He considers an establishment of social habits of transferring debt as equivalent to transferring money as a necessary historical foundation for development of banking (Keynes 1971 [1930], p. 20).

economic position and gives them power with respect to the rest of the economy (Keynes 1937a, p. 666, 1937b, p. 248). In a sense, control over liquidity is equivalent to control over investment and, more broadly, economic activity.

According to Keynes, the pool of liquid resources is constituted by existing cash and the new cash created by banks (Keynes 1937a, p. 665). Put differently, banks perform their pivotal function of liquidity provision by issuing claims against themselves. They do it in two ways – by intermediating existing liquidity and by creating new (Keynes 1971 [1930], p. 21-22). The former corresponds to their passively collecting deposits, the latter – to their actively lending out claims on themselves.^{33 34} On these grounds Keynes treats banking business as a business of creating and cancelling deposits and balancing active and passive deposits. He also stresses that the two types of liquidity are equivalent not only from the banks' perspective, but also from the borrower's perspective (Keynes 1937a, p. 664-5, 1937b, p. 245-6, 1938, p. 319).³⁵

This take on banks as both active and passive has important implications. First, banks for Keynes are both intermediaries and money creators, and there is no dichotomy between the two aspects of the banking business. Banks do perform an intermediation

³³ Keynes draws on a similar argument by Phillips (1920, p. 40) who distinguished between a primary and a derivative deposit. The former “arises from the actual lodgment in a bank of cash or its readily convertible equivalent”, whereas the latter “arises directly from a loan” in anticipation of its repayment.

³⁴ Although Keynes is right in distinguishing between these two functions, it is not quite correct to associate deposit collection with a passive function. Both deposit taking and issuing claims against itself in favor of a borrower involve active banking aiming at increasing the lending capacity of banks. The active principle behind deposit collection by means of offering interest on deposits and making them means of circulation was rightly stressed by Hicks (1989, p. 56-57).

³⁵ This is further developed by Fischer and Rossi. The latter emphasizes the co-existence of two functions of banking: “money-purveying”, i.e. creation of new liquidity, and “credit-purveying”, i.e. intermediation of existing liquidity (Rossi 2003, p. 340, 347-348, based on Rossi 1998). Therefore, banks for Rossi are both monetary and financial intermediaries, hence, there exists a two way causality: loans make deposits, and deposits make loans (Rossi 2003, p. 350-351).

function, but it is only one aspect of their business, therefore they cannot be theorized as primarily intermediaries. By the same token, banks cannot be viewed as essentially money creators, although they do create money, henceforth it is misleading to argue that they are not intermediaries. Keynes' focus on liquidity provision gives a framework to nest both elements. As a result, Keynes goes beyond Schumpeter whose theory emphasizing money creation *ex nihilo* can be viewed as a particular case of Keynes' liquidity provision. It is not surprising, because, as was argued above, Schumpeter's emphasis was largely driven by his focus on development, which can shed light only on a particular type of banking.

Second, Keynes' emphasis on these two modes of performing banking business is important for locating his insight that loans create deposits and that liquidity is distinct from saving. Keynes clearly distinguishes between two questions concerning deposits – their being a limit on liquidity and their role in banking activities. Although deposits indeed do not set a limit on the bank lending capacity, and saving – on liquidity, it does not make deposit collection irrelevant for the banking business. This is wholly in line with a casual observation that even modern banks engaging in a vast array of activities ranging from asset sales to trading securities do not cease to collect deposits. Therefore, Keynes' distinction between the question of limits on loans and liquidity and relevance of particular ways in which loans are funded needs to be retained.

Third, Keynes centres his discussion of liquidity provision around the finance motive for holding money, i.e. around the investment needs of enterprises (Keynes 1937b, p. 246, 1939, p. 574). Thus, contrary to Schumpeter, Keynes does not reduce

banking activities to financing development, but treats them broader as providing finance to enterprise in general.

Finally, although Keynes does acknowledge the money creation function of banks, he also stresses that this process has its limits:

I have endeavoured to say enough to show that the familiar controversy as to how and by whom bank deposits are ‘created’ is a somewhat unreal one. There can be no doubt that, in the most convenient use of language, all deposits are ‘created’ by the bank holding them. It is certainly not the case that the banks are limited to that kind of deposit, for the creation of which it is necessary that depositors should come on their own initiative bringing cash or cheques. But it is equally clear that the rate at which an individual bank creates deposits on its own initiative is subject to certain rules and limitations. (Keynes 1971 [1930], p. 26)

Some of the bank obligations, including inter-bank liabilities, are settled in cash, calling for bank reserves. This holds prior to and regardless of the government regulations requiring banks to keep a certain amount of reserves. Therefore, even if banks were to “move forward in step” (Keynes 1971 [1930], p. 23) – lend out claims on themselves in favor of the borrowers at the same pace as other banks – there would still be a limit to credit. This insight is important, because it reveals that Schumpeter’s abstract process of money creation has its limits. What is more important, it shows that the existence of many banks places constraints on the credit creation, therefore, paraphrasing Marx (1977 [1894], p. 250), the real barrier of the banking credit becomes capitalist banking itself. The banking system spontaneously emerges as a means to overcome the limits of commercial credit, yet the capitalist form of banking can never achieve the goal of removing limits of credit expansion.³⁶

³⁶ This has two implications. First, the benefits of credit expansion could be more easily realized if banking were placed on a non-capitalist basis. It could have eliminated the limits to credit associated with the functioning of the capitalist credit system itself, and would have shifted the limits to where they actually belong – the sphere of conscious planning of economic activity. Second, it gives an insight into the tendency towards the rising concentration in banking. Individual banks strive to overcome the limits

Keynes' insights are important for the purposes of the present study for a number of reasons. First, liquidity provision can be viewed as a trans-historical function of banking that has not been challenged by its ongoing transformation. The concrete ways in which banks provide liquidity have evolved, for instance, commercial and industrial loans have often been replaced by contingent liabilities – banks promises to provide liquidity on demand, but the underlying activity of liquidity provision has remained intact. As was argued in chapter 2, this cannot be said of the information collection function, as in many cases it has been reduced to credit scoring, outsourced to rating agencies, and mitigated by rising availability of financial information with a development of the information-telecommunication technologies. If a constant bank profit share (and a rising profit share of all financial institutions) does not sit comfortably with the decline in the information function, it can be partly explained by the liquidity provision function that has not gone away. Therefore, it is not surprising that one of the new approaches to banking in the mainstream literature pivots on banks as essentially liquidity providers. In a sense, Rajan's (1996) theorization of banks as essentially liquidity providers on demand, both on the asset and the liability sides of their balance sheets, is rooted in Keynes' insight, although it is not acknowledged by Rajan.

Second, Keynes' emphasis on banking as liquidity provision clearly shows why banking should not be conflated with the intermediation function as its particular type. This in turn implies that a decline in the relative importance of the traditional banking

associated with reserves, and accomplish it to the extent that their market share increases. A similar argument was advanced by Bossone (2001a, p. 872, 2001b, p. 2254, 2264) and Graziani (2003, p. 91). It is important to stress that contrary to the possibility of socialization of benefits associated with credit expansion on a non-capitalist basis, the capitalist solution involves banks privately appropriating these benefits.

represents a transformation, but not a decline in banking. Within Keynes' framework, it is not surprising that the mainstream empirical literature discussed in chapter 2 has evolved from a discussion on the decline in to the transformation of banking.

In spite of these two major insights relevant for understanding banking and its transformation, there is an important limitation of Keynes' treatment of banks. The emphasis on liquidity provision is not a sufficient foundation for theorizing banks. Keynes is right that banks are a monetary phenomenon, nevertheless, they are more than that – they are a phenomenon of credit. The latter is more complex and has more determinants. A credit relationship can be understood as an exchange of money for a promise to pay – an act involving two moments. One moment of this transaction is indeed liquidity provision. Nevertheless, the other aspect is an extension of a promise to pay. Even though these are two sides of the same coin, Keynes' emphasis on liquidity provision needs to be shifted to that on the promise to pay, as the latter involves more determinants. This is clear from Marx's analysis of functions of money, in particular, money as a universal equivalent (Marx 2003 [1867], p. 74-75) and money as a means of payment (Marx 2003 [1867], p. 134-141). At the most elementary level, liquidity provision is grounded in the function of money as coin – universal equivalent. A need for liquidity arises because money is in a monopolist position of exchangeability with all the other commodities. Therefore, liquidity provision is a provision of universal equivalent. By contrast, issuing promises to pay is based on the function of money as money – means of payment. It subsumes the function of universal equivalent as one of its aspects, but money as a means of payment also implies further social relations that cannot be analyzed at the level of money as universal equivalent. Among them are economic factors

related to a separation of purchase and sale, and non-economic determinants, especially, trust between the counterparties involved in a credit relation. Therefore, a shift in emphasis from liquidity provision to credit essentially implies broadening the scope of analysis and creating a room for a discussion of aspects that cannot be captured merely on the basis of liquidity. A possible way of doing this will be discussed in chapter 4.

To recap, Keynes stays at the roots of the modern heterodox theories of banking, in particular, of the circuitist and post-Keynesian approaches. Liquidity provision can be viewed as a trans-historical function of banking not challenged by its ongoing transformation, therefore, can serve as a broad foundation for theorizing banking. In spite of this strength, Keynes' focus on liquidity should be supplemented by an accent on an exchange of promises to pay, as it could give more mileage for uncovering social relations underlying banking activities.

3.3. Theory of the Monetary Circuit: An Approach to Banking and Its Transformation

An approach to banking within the monetary theory of production has emerged as a synthesis of Schumpeter's and Keynes' views on banking. The proponents of the theory of the monetary circuit, also known as the circulation approach, adopt the general Schumpeterian framework by considering creation of purchasing power as the fundamental function of banking. But they also alter Schumpeter's approach by shifting the emphasis away from credit for development to credit to start any circular flow, be it a new flow related to innovation or a new circular flow within the already established enterprise (what Schumpeter calls "current credit"). They consequently broaden the scope

of the analysis of banking. In this shift of the focus the circuitists follow Keynes' (1937a, 1937b, 1938, 1939) discussion of the "finance motive for holding money".

Therefore, contrary to the dominant mainstream theories discussed in the previous chapter, the circuitist theory places banks on a foundation other than information asymmetries and market imperfections. At the same time, by acknowledging differences across economic sectors and explicitly positing banks in their relationship with these sectors, the monetary theory of production overcomes the limits of both the old and the new mainstream theories of banking. On these grounds, this theory appears well suited for understanding the ongoing transformation of banking and bank profits. Consider the core of the theory in a greater detail to show to what extent it is indeed a good framework for studying this transformation.

3.3.1. Banking in the Theory of the Monetary Circuit

The starting point of analysis of the circulation approach is an observation that the process of production requires initial finance to pay the wage bill in advance.

Aggregation of all the firms into one sector allows the circuitists to abstract from payments for means of production, as they cancel out within the firm sector. In the absence of government, and a past stock of wealth in a money form, the only source of purchasing power can be banks that create money *ex nihilo*.

Banks are thus derived as specialized institutions with two main characteristics. First, they provide initial finance to start a circuit (Graziani 2003, p. 27, Bossone 2003, p. 154, Bossone 2000, p. 25, 34). Sometimes the state is also included in the analysis as the other agent "whose expenditures are creating present and future real wealth" (see also

Parguez 2004, p. 20, Parguez 2001, p. 72, Parguez, Seccareccia 2000, p. 105). The relevant emphasis here is that from the outset lending is considered only for the purposes of wealth creation, therefore, other types of lending are left out. Parguez and Seccareccia (2000, p. 107) justify this by saying that issuing liabilities to finance non-wealth generating activities would make these debts “deprived of value and no one would accept them”³⁷. Second, banks provide initial finance by creating money ex nihilo, or lending out debts on themselves (Bossone 2001b, p. 2240, 2247, 2252, 2266, Graziani 2003, p. 25, 2004, p. 26, Parguez 2001, p. 73). This is grounded in an approach to money as necessarily credit money. Parguez (2002, p. 46) summarizes this idea by saying that a specialized institution becomes a bank when it is viewed as “so credit worthy that their debts are universally accepted as means of acquisition” and they start specializing in issuing debts on themselves.

This take on banking has three important implications. First, money creation is independent of the multiplier process which is merely a multiplication mechanism, but not the origin and the simplest form of money creation (Bossone 2001b, p. 2250). Second, for the circuitists, there is no pool of pre-existing deposits for the banking sector as a whole, therefore, banks as a group are not intermediaries (Parguez 2001, p. 79, Bossone 2003, p. 154). Parguez and Seccareccia (2000, p. 107) summarize it arguing that “loans can never be financed by some pre-existing deposits”. Third, this ability to create money is a source of banks’ power (Bossone 2001a, p. 858, Graziani 2003, p. 150). On

³⁷ This justification is obviously problematic for both historical and logical reasons. Lending for a variety of purposes has always been a really existing phenomenon. By the same token, borrowing for purposes other than wealth generation can co-exist with these debts being credible, when the debt is expected to be repaid from other sources of revenue accruing to the debtor.

these grounds Parguez (2004, p. 32) views banks as being able to set interest rates at their discretion³⁸.

The circuitist approach to banking contains several relevant insights that give this theory an advantage over the mainstream formulations. The proponents of the monetary theory of production distinguish among economic sectors and locate banks among them at the outset. By the same token, they rightly treat banks as a monetary and not an information-collection or a risk-related phenomenon. Finally, they correctly emphasize that banks cannot be viewed as passive intermediaries. Nevertheless, the exact way in which the circuitists develop these relevant points is often problematic.

First, banks are indeed in a unique position to create money by lending out claims against themselves. But this is not a good starting point for an analysis of banking, as it cannot serve a basis for a discussion of other bank activities. The circuitists are right in modifying Schumpeter's perspective by shifting the emphasis from development to initial finance. Nevertheless, they have not realized that the two aspects of Schumpeter's theory of banking – the development lens and money ex nihilo – are intertwined. As was argued above, in Schumpeter money needs to be created ex nihilo because before the process of development starts there can be no savings from the past and because the already existing liquidity is assumed to be channeled into the established circuits. This take was shown to be problematic, as there is no need to assume away existing monetary wealth even in the context of development. Nevertheless, acknowledging that development is not a necessary condition for banking only reinforces the fact that there is no need and, in fact, no theoretical basis to consider credit as necessarily based on new purchasing power.

³⁸ The same emphasis on banks as price setters, i.e. interest rate-setters, can be found in some post-Keynesians, for instance, Moore (1988, p. 57).

That is why Keynes (1937a, p. 664-665, 1937b, p. 245-246, 1938, p. 319) is right arguing that the “finance motive” can be satisfied by either existing or new liquidity, therefore, the need to view the latter as the only source of liquidity disappears. This, in turn, discards the notion of banking as fundamentally a creation of new purchasing power, advocated by Schumpeter and later by the circuitists.

The circuitists’ emphasis on banking as money creation reveals a conflation of two separate questions – that of the content of banking activities and of the character of modern money. The fact that money nowadays is credit money created by banks – a point rightly stressed by the circuitists – does not imply that banking is essentially an activity of money creation. This conflation occurs because banks are derived by the circuitists at a wrong level of abstraction – that of money, instead of a specific capitalist enterprise corresponding to a particular circuit of capital. Not only the logic, but also the concrete reality of banks performing the intermediation function requires this element of banking to be retained.

When it comes to the question of how banks operate, and not how money is created, following Keynes, it is important to distinguish between two questions concerning deposits. The one is whether deposits constitute a limiting factor in liquidity provision, the other – whether they are essential for the content of banking. While Keynes has indeed argued that deposits do not represent a limit on liquidity provision, it does not imply that they are irrelevant for the banking business. Hicks holds a similar perspective, but he strengthens it by effectively shifting the focus away from the question of what is prior – deposits or loans – to the logic of development of the functions of banking. He aims at establishing the simplest form of banking activities, from which all the others can

be derived as its particular forms. This approach is in line with the take on abstraction suggested above. Graziani is aware of Hicks' work, but he seems to fail to understand that Hicks deliberately shifted the focus, as he argues that "it would be hard to say on which side he [Hicks] stands and whether he assigns a priority to loans or to deposits" (Graziani 2003, p. 82).³⁹ By the same token, Graziani fails to understand the importance of Hicks' perspective.⁴⁰

Thus, banks indeed can create purchasing power, but it is one of the forms of liquidity provision by banks, therefore, the circuitist theory of banking should be viewed as a partial case of a broader theory of banks as liquidity providers. This take on banking was originally suggested by Keynes (1971 [1930]) and has recently reemerged in the mainstream literature in the context of the transformation of banking (Rajan 1996).

Second, the circuitists are right in their emphasis on the need to posit banks in their relationship to other sectors of the economy, nevertheless, the specific way they do it is limited. Their theory of banking fits the realities of traditional banking as it existed in the 20th century, when banks were the major liquidity provider to firms. However, this cannot be a basis for an analysis of banking in general, as this take on the content of banking is too narrow: credit can be extended for purposes other than starting a circuit, on the one hand, and a circuit can be started without credit, on the other hand. The fact that in a certain historical period banks' primary business was lending to non-financial

³⁹ Other circuitists are also aware of Hicks' emphasis on discounting as the origin of banking and deposit taking as a prerequisite of money creation (Bossone 2001b, p. 2244), but do not see what it implies for their theory.

⁴⁰ Circuitists usually deny the logical priority of demand deposits over issuing notes stressed by Hicks. Parguez goes further by denying even moneyness of demand deposits: "Hoarded deposits are not money but mere financial assets. Money cannot be hoarded in the modern capitalist economy where it is banks money" (Parguez 2004, p. 32).

businesses does not mean that this should be a foundation of a general theory of banking. This in turn implies that the circuitists' theory of banking is a partial theory of banking, stressing relevant activities and rightly emphasizing the difference between firms and workers, yet failing to provide a foundation for a broader theory of banking. It is not surprising, given that they do not seem to aim at doing it, in the first place, and their primary concern is the theory of production and its monetary aspects, but not a theory of banking per se.

3.3.2. The Circuitist Theory of Banking in the Context of the Transformation of Banking

The problematic character of the circuitist theory of banking becomes apparent in the course of the transformation of banking⁴¹. It involves both constituents of this take on banks.

First, money creation ex nihilo does not sit comfortable with the ongoing transformation of banking, for a number of reasons. On the one hand, the deposit taking function of banks has not disappeared, and a general theory of banking should allow for this. A casual observation shows that collecting deposits is one of the really existing banking activities, which can be understood as centralization of funds scattered across the society that gives money capital banks deal in a truly social foundation. Incidentally, looking at deposits collection from this angle gives a deeper meaning to this seemingly simple activity, which makes retaining the aspect of intermediation even more important.

⁴¹ This does not concern all the aspects of their take on the transformation of banking, as the circuitists have some interesting insights into e-money and, to some extent, a rise in bank concentration (see, for instance, Bossone 2001b).

On the other hand, banks have been increasingly involved in asset sales and securitization. If banks can create money ex nihilo, it is difficult to explain why loans would evolve into marketable securities, i.e. why the process of securitization would take place. Within the circuitist approach, banks could have just issued money ex nihilo to further expand their lending business. By the same token, institutions within the shadow banking system – say, special purpose vehicles or government sponsored enterprises – do not create money. Therefore, the circuitist theory cannot help explain their business. This naturally leads to treating it as speculation – a widely deployed approach, which nevertheless does not give much mileage. Finally, an emphasis on money creation ex nihilo makes the question of how banks fund themselves appear irrelevant. Nevertheless, as the actual banking practices show it is a relevant matter for banks. Moreover, the mode of funding can have important implications for the nature of bank revenues, as it will be shown in chapter 5. Therefore, Keynes’ emphasis on liquidity, both existing and new, discussed in the previous section can offer more insights and needs to be retained.

The other constituent of the circuitist theory of banking – lending to provide initial finance – also does not stand the test of the transformation of banking. A rise in lending to households and non-lending bank activities has made it clear that if the essence of banking business is defined through lending for productive purposes, modern financial institutions are not banks. In this sense, this take on banking would be another theory of disintermediation, which cannot be reconciled with a rise in the profit share accruing to banks, as was discussed above in the context of the mainstream theories. To consider lending to agents other than firms as banking, one would need to either step back from

the idea of starting the circuit as an essence of banking, or try to redefine what starting a circuit means. It is the latter approach that has been adopted by the circuitists.

Trying to preserve the conceptual core of the theory, Graziani suggests treating lending to households as indirect lending to firms that allows consumers to buy finished products from these firms (Graziani 2003, p. 21). Bossone has a similar argument with respect to rising bank lending to other non-bank financial institutions (Bossone 2001b, p. 2261). They do not seem to recognize that although by interpreting lending to households and other financial institutions as an indirect lending to firms they can shift the emphasis back to bank lending for production, it nevertheless does not preserve the core of the theory. It happens because this attempt also involves shifting the focus away from starting a circuit of production to completing the circuit, which is still different from the way the circuitists define the key function of banks. In addition, it is not clear what are the insights offered by this interpretation.⁴²

Parguez has a more sophisticated take on lending to households. It starts very similarly to the argument proposed by Graziani and Bossone when this type of lending is interpreted as “an increase in the roundaboutness of the monetary structure” (Parguez 2004, p. 46), when instead of directly creating money for firms, banks channel money to them indirectly. There are two types of intermediations for Parguez – loans to wage-earners for consumption and acquisition of stocks. The former generates “excess profits” for firms (over and above investment), either through increased spending by households or through their buying stocks resulting in increased investment by firms. The argument

⁴² In general, this attempt to interpret lending to households as an indirect lending to firms resembles other similar instances in the history of thought when additional ad hoc assumptions would be introduced into a theory failing to account for new evidence or changing circumstances. Ptolemaic astronomy and Newtonian mechanics challenged, respectively, by Copernicus and Einstein, are classical examples of it.

about rising roundaboutness is further used by Parguez to point out its unsustainability, thus, he concludes a crisis would restore a monetary structure with a lower roundaboutness. The argument by Parguez gains a substantial analytical advantage over Graziani's and Bossone's, when the former also explicitly connects these developments with bank profit. He is clear that rising lending to households has become the main source of accumulation by banks (Parguez 2004, p. 47) that in addition shifts the interest burden from firms to wage-earners. This develops his earlier argument that households' net new debt "reflects the net contribution of households to profits" (Parguez 2002, p. 50) – a point that was left unelaborated in his earlier work.

As a result, with a partial exception of Parguez who at least acknowledges the different content of bank profit coming from lending households, the way the circuitists conceptualize the link between the circuit and the banks becomes self-defeating. The proponents of the theory of the monetary circuit start by acknowledging the difference across economic sectors, which is one of the main strengths of their take on banking over the mainstream theories. But in their treatment of the transformation of banking they end up abandoning this difference across the sectors, therefore, they lose their own advantage over the mainstream theories of banking and do not realize the potential of the theory embedded in its original set up.

From the perspective of the present study, the general framework of the monetary theory of production, with its emphasis on differences across economic sectors, puts the theory in a good position to analyze the transformation of banking, if it were not to limit itself to only lending for production as the content of banking activities. This is where the circuitists could have used another insight of Schumpeter on whom they draw heavily. In

his criticism of Aristotle, Schumpeter was very clear that interest from lending for production and for consumption has a qualitatively different nature. He was critical of Aristotle who “did not even classify loans according to the various purposes they are capable of serving and does not seem to have noticed that a loan that financed consumption is something very different from a loan that financed maritime trade” (Schumpeter 1954, p. 65)⁴³. Therefore, insights into the nature of interest from lending for production cannot be directly applied to interest from consumer borrowing. This idea is in general neglected by the circuitists who try to reinterpret lending to households in a way to fit it into their core idea of productive lending as the main business of banking. The only exception is Parguez who to some extent goes beyond this view, but does not elaborate on it.

Incidentally, the rising household indebtedness has revealed not only limits of the theoretical framework of the circulation approach, but also a fundamentally problematic nature of Graziani’s class analysis. For him, only firms have an access to credit, hence, unlimited purchasing power, whereas wage earners are excluded from credit and can spend only the already earned income (Graziani 2003, p. 19-26, 98, 145-146). This approach has not stood the test of time, and the ongoing changes in banking have shown that an access to credit cannot be a basis for class distinction, a criterion of one’s class position. On the contrary, it is the class position that affects the terms and the content of

⁴³ An idea that interest can have a different content depending on the purposes of borrowing can be found elsewhere in Schumpeter (1954, p. 327, 647), although he never elaborates on it. For a more recent appreciation of the differences between types of credit see Turner (2010, p. 42) who argues that “whether this increase [of bank balance sheets and leverage in the real economy] was value added depends crucially on the economic and social functions which credit performs, that these functions vary by category of credit”.

the credit relations one is involved in. Thus, the causality runs from the class position to the specificity of credit relations, not from access of credit defining one's class position.

Finally, the transformation of banking reveals weaknesses of the monetary theory of the circuitists. Parguez argues that "if new money were not instantaneously spent to generate real value, it would be deprived of any extrinsic value" and that the nature of money lies in its being "a means of acquisition of real resources in order to generate more future real wealth" (Parguez 2001, p. 80). This approach is obviously problematic, and Parguez immediately tries to reverse the argument by arguing that in the case of banks extending credit to finance acquisitions of financial assets "money is still created" (Parguez 2001, p. 80). It is obviously hard to reconcile these two statements.

To recap, the circuitist theory of banking offers relevant insights that need to be retained, namely, an emphasis on differences among economic sectors and a monetary nature of banking activities. Nevertheless, the exact way these two features are developed in the monetary theory of production encounters its limits in the course of the transformation of banking and calls for a broader foundation for theorizing banks. To some extent, it is not surprising that the theory of the monetary circuit fails to provide a self-standing theory of banking, because in this theory banks appear at a wrong level of abstraction. They emerge in the context of money, derived through money (Graziani 2003, p. 17, 152). Instead, they should be analyzed in the first place as specific capitalist institutions performing a certain set of functions and driven by a profit motive.

3.4. Bank Profit in the Theory of the Monetary Circuit: Interest as Rent

The approach to banking within the monetary theory of production gives a rise to a peculiar theory of bank profit as rent. This in turn serves a basis for treating banks as rentiers (Parguez 2004, p. 26-44, 2006, p. 140) or even the leading arm of the rentiers class (Parguez 2004, p. 32, 41), because their sole source of income is interest (Parguez 2004, p. 33).

The circuitists usually treat interest accruing to banks as rent, or seigniorage, on one of the two grounds (or some combination of the two). Bank profit is either rent in the sense of coming from a pure redistribution of a given amount of surplus based on a conflict of interest between a lender and a borrower, or a quasi-monopoly rent accruing to banks due to their exclusive power as money creators that do not bear opportunity costs of postponed consumption. Consider both lines of reasoning in a greater detail.

The first approach to interest accruing to banks as rent pivots on the redistributive character of this form of revenue. Banks do not create value, but participate in distribution of value created by the firm sector (Graziani 2003, p. 150, Parguez 2004, p. 39, Graziani 1987, p. 31, Parguez 1987, p. 112-113). A similar line of reasoning can be found in Marx's (1977 [1894], p. 372-379) discussion of the division of surplus value into interest and profit of enterprise⁴⁴. A formally similar take is also advanced by Schumpeter for whom, as discussed above, interest is a kind of tax on entrepreneurial profit (Schumpeter 1961 [1911/1934], p. 175, 210) that involves a "price struggle" between the bank and the entrepreneur (Schumpeter 1961 [1911/1934], p. 125, 192-193). The proponents of the monetary theory of production accept the idea of the redistributive

⁴⁴ The direct relation to Marx's argument is explicitly acknowledged by Lunghini and Bianchi (2004, p. 158).

character of bank profit, and interpret this redistribution as rent extraction. An approach to interest as rent is most clearly stated in Parguez, for whom interest is “a pure exogenous constraint on firms” (Parguez 2004, p. 27, see also p. 40) that controls the distribution of net wealth between firms and rentiers. He views this redistribution as an “income drain” (Parguez 2004, p. 33)⁴⁵. This argument takes a particular form in Lunghini and Bianchi, for whom in the processes of production, income distribution, and capital accumulation (and class conflict), interest plays the same role as rent, and banks as landlords in Ricardo’s theory of distribution (Lunghini, Bianchi 2004, p. 153, 157-158). This happens because the rate of interest does not participate in relative price determination, so that the division of surplus value is a transformation of “unpaid labor into profit and financial rent” (Lunghini, Bianchi 2004, p. 159). This is equivalent to arguing that interest is a pure parameter of redistribution, and any redistribution of a given amount of surplus is equivalent to rent extraction.

Although the circuitists are right in their emphasis that no value is created in the sphere of circulation, hence, that interest originates in redistribution, it is misleading to view interest as rent on these grounds. It is important to bear in mind that there is a variety of forms of income based on redistribution of the already created value, with each of them having different underlying social relations and concrete mechanisms of profit extraction. Put differently, not every form of income coming from redistribution of flows of value is rent. Marx was clear about it emphasizing that in general profit of enterprise, commercial profit, interest, and rent all stem from redistribution of the already created

⁴⁵ For Parguez “interest payments are part of production costs, they are financed by money creation. Banks can play freely their part of rentiers because they have the power to advance to firms the amount of their interest income” (Parguez 2004, p. 26-27). This emphasis on money creation by banks in relationship to their profits is connected to the second take on bank profit as rent that will be discussed below.

surplus value (Marx 1977 [1894], chap. 17, 19, 22-23, 37, 48). Yet, in each of the cases redistribution is grounded in different social relations. For him, an object of analysis should be the content of these relations that give a stream of income a specific economic character and justify its existence, and not an attempt to establish formal similarities among the forms of revenue originating in the sphere of circulation. There are two important aspects of it.

First, among the other reasons, the emphasis on the social relations attached to certain redistributive outcomes is important because it shows that they often attain rationalization and some sort of social justification. This is the main difference between Ricardo and Marx. Harvey (2006, p. 331-332) stresses that, contrary to Ricardo who considered rent as a parasitic form of income, Marx's goal was to show how and why rent extraction obtains social justification and, therefore, its analysis cannot start with an argument about its parasitism. Put differently, for Marx, the real problem is not that landlords are parasites, but why and how they become a necessary element of the capitalist mode of production. To rephrase it, the real problem is not that banks are parasites, but how and why their activities – even the most destructive ones – become necessary in capitalism.

Second, an important implication of this emphasis on the differences across forms of redistribution is that interest is not rent. On the capitalist basis, interest develops as a price of a sum of value loaned out as capital (Marx 1977 [1894], p. 345).⁴⁶ By contrast, rent is “the specific economic expression of landed property” (Marx 1977 [1894], p. 622,

⁴⁶ Itoh and Lapavitsas (1999, p. 65) argue that Marx is wrong in his understanding of interest that should rather be seen as a reward for parting with a sum of value that should not necessarily be loaned as capital – it should simply be loaned.

see also p. 618, 624)⁴⁷. Therefore, the origin of interest can be traced back to the special social status of money as universal equivalent, whereas the origin of rent is rather linked to property rights. Incidentally, the circuitist' confusion between interest and rent is not unprecedented in the history of economic thought (Marx 1977 [1894], p. 622-624). Marx was critical of Carey who treating ground-rent as identical with interest has overlooked the specific character of rent. The opposite mistake was made by Dudley North, Locke, and Turgot when they viewed interest on capital as a form analogous to rent, thus, overlooking specific social relations underling interest payments.

Incidentally, the conflation between interest accruing to banks and rent might involve further misunderstandings. For Marx, property rights allow a rentier to exclude someone from an access to a resource (say, land), hence, determine the economic content of rent as a form of revenue. Although banks do have a power to constrain liquidity, clearly, this power does not hinge on property rights, as banks do not own loanable capital they operate with – they are rather general managers of loanable capital of the society as a whole (Marx 1977 [1894], p. 368, 402). Therefore, considering interest accruing to them as rent would be even more misleading than in the case of interest accruing to the owners of the temporarily idle money. Nevertheless, a specific take on banking within the monetary theory of production might provide a parallel between

⁴⁷ Another interpretation of the social relations underlying rent extraction was developed by Teixeira and Rotta (2010). For them, the economic character of rent in Marx's theory is determined by lending out use-values, as opposed to lending out a sum of values corresponding to interest-bearing relations (Teixeira, Rotta 2010, p. 23). This is grounded in Marx's argument that "the price of things which have in themselves no value... may be determined by many fortuitous combinations" (Marx 1977 [1894], p. 633, see also p. 817). Therefore, a monopolized access to this type of things opens a room for rent extraction. Although monopoly rights play an important role in the argument by Teixeira and Rotta, nevertheless, it is primarily driven by what is being lent out – use value or exchange value. Bearing that in mind, clearly, bank lending is obviously lending of a sum of value, hence, cannot be theorized as rent extraction, in spite of their monopoly over money creation.

interest and rent precisely for the reason that banks are treated as creators of money ex nihilo. The circuitists cannot recognize the difference between rent as an economic form of property rights and interest as an economic form of the special status of money as universal equivalent, because for the circuitists banks lend out what they create, and in this sense can be viewed as owners of money. This might be used as a justification to treat interest as rent, which brings us to the other approach to interest as rent within the monetary theory of production.

The second take on interest accruing to banks as rent can be found in Graziani, Parguez, and Bossone. For them, interest is a quasi-monopoly rent accruing to banks due to their privileged position in money creation. This privileged position gives them power⁴⁸ in their relationship with firms, which in turn serves a basis for “a transfer of real wealth from industry to finance” (Graziani 2003, p. 103). He makes an even stronger statement that interest on bank loans is “the only real costs” to firms, while “monetary payments made to wage earners are never a real cost to firms” (Graziani 2003, p. 116). Bossone further develops this argument . For him, given that banks are the only capitalist institutions capable of money creation ex nihilo and given that they have no opportunity costs in the form of forgone consumption and do not provide intermediation services, their profit is quasi-monopoly rent (Bossone 2001a, p. 871, 880-881, 884, Bossone 2001b, p. 2243, 2253, 2266, Bossone 2003, p. 155-158, Bossone 2000, p. 25-26).⁴⁹

⁴⁸ The emphasis on power can be traced back to Keynes, for whom the source of banks’ power is their control over liquidity, as was argued above. The difference between Keynes and the circuitists is in the fact that banks’ power has a broader foundation in Keynes. It is control over liquidity, with creation of new liquidity being its particular form.

⁴⁹ This approach is directly related to an argument that banks are quantity-takers and price-setters that have a power to set an interest rate to accommodate their desired rate of accumulation, allowing them to achieve “the forced redistribution of wealth between firms (productive capitalist class) and banks (impersonating

From the discussion above it is clear that this approach to bank profit is misleading because it is an immediate consequence of the circuitist theory of banks as essentially money creators. If banks are considered as primarily creators of purchasing power *ex nihilo*, it trivially follows that there are no opportunity costs involved, hence, that interest they receive is a pure rent. In this sense it is not surprising that the circuitists reach a conclusion that interest is rent. Nevertheless, as was argued before, this is a limited take on banking because although banks do have a capacity of lending out their own debt, they also intermediate existing liquidity. Therefore, the general character of bank profit as rent cannot be derived from a partial take on banking focusing on its particular activity. Furthermore, the economic character of bank profit cannot come from the concrete mechanism of generating this liquidity. Whether banks create own money or use deposits is a matter of secondary importance for the content of bank profit that rather stems from bank activities and their relationship to the circuit of capital and revenues.

Incidentally, this becomes apparent in Bossone's own treatment of bank profit (Bossone 2001a, p. 871, 2001b, p. 2253, 2264, 2003, p. 156). He starts his analysis with identifying interest accruing to banks with quasi-monopoly rent and interest received by non-bank financial intermediaries not with rent, but profit for their intermediations services, given that there is no process of liquidity creation involved. The problem is that the same holds for banks as well – to the extent that they also intermediate liquidity. Therefore a line between returns on new liquidity and on existing liquidity cannot be drawn between banks and non-bank financial institutions, respectively, and another approach to bank profit should be adopted.

the rentiers class)" (Parguez 2004, p. 27). By raising interest rates, "rentiers increase squeeze on firms" (Parguez 2004, p. 39). An argument about banks as price setters was also advanced by Moore (1988, p. 57).

A corollary to this proposition is that interest cannot be treated as rent, because there exist costs of bank lending associated with, first, interest on deposits, second, costs of fixed and variable capital and, third, costs of reserves. If the first form of expenses is related to the intermediation business of banks, the last two are incurred even in the case of banks' lending out claims on themselves.

This is not to deny that there is a grain of truth to the argument about the monopolistic position of banks in their ability to create money. It is used by the proponents of the circulation approach as an argument that all bank profit is quasi-monopoly rent, but a more interesting question from the perspective of the present study is why banks come to hold the monopolist position of money creation and how this monopolist position is sustained by the social relations of trust that is created and generalized through functioning of the financial system itself (Lapavitsas 2003). Put differently, banks monopolistic control over liquidity is not accidental and is rather a form taken by the necessarily social content of the business of liquidity provision.

To recap, first, the monetary theory of production approach to bank profit rightly stresses that this profit originates in the sphere of circulation and has a redistributive character. Banks do not create value, but receive a part of value created elsewhere. Nevertheless, it is misleading to conclude on these grounds that bank profit is rent. This kind of generalizations based on formal similarities among different forms of income redistribution should be avoided. The concern should be shifted to uncovering social relations underlying different forms of income distribution, and these relations in the case of interest payments are different from those in the case of rent. Therefore, bank profit is,

first, not rent which is an economic form of property relations and, second, far from being homogeneous.

Moreover, second, conflating rent with interest runs a risk of further confusions. Contrary to rent, interest received by banks is not based on property relations, as banks do not own loanable capital they operate with. This problem obviously does not arise for the circuitists because in their theory banks produce money and in that sense can be said to own what they lend out. By the same token, the concept of rent carries over a connotation of an idle, parasitic rentier. References to Ricardo reinforce this, as he treated landlords as a parasitic stratum. When it comes to banks, they cannot be theorized along these lines, although there might be an element of parasitism in their activities.

Finally, the economic character of bank profit comes not from the general function of banks as liquidity provides, and even less from a particular form of liquidity provision based on creation of new purchasing power, but from activities banks engage in and from resulting links to flows of value.

These three sets of conclusions become even more important in the context of the developments over the past several decades. The take on bank profit as rent becomes increasingly problematic, because it mistakenly associates the rise in the bank profit share with a quantitative, instead of a qualitative change. For example, some proponents of the monetary theory of production interpret the rise as a rise of “a pure rentiers economy” which is the late stage of the capitalist economy (Parguez 2004, p. 41, Parguez 1987, p. 115). Others look at it as a “revenge of the rentier” associated with a rise in rentier income to the detriment of non-rentier profit (Rochon, Rossi 2010, p. 7-8). Yet others understand it as a rise in the seigniorage power of banks (Bossone 2001a, p. 884).

From the perspective of the present study these approaches, instead of shedding the light on the transformation of banking, divert the attention from and become an obstacle for posing an important question – that of the changing composition of bank profits. When bank profit is viewed as rent, neither new sources of bank profit, nor changes in the balance between the established sources of profit can be recognized. Therefore, this approach does not allow one to recognize that the ongoing rise in bank profits is not a quantitative increase in bank profits coming from a rise in their monopoly power over liquidity creation, but rather a qualitative change associated with a changing nature of their activities reflected in an altered composition of bank profits. Put differently, the rise in the bank profit share should be associated with a structural change in banking, and the argument by DeYoung and Rice (2004a) that the changing composition of bank revenues is a useful lens for studying the transformation of banking should be retained.

3.5. Financial Profit in the Literature on Financialization

It was argued in the previous section that in spite of some strengths of the approach to banking within the monetary theory of production, this theory does not give much mileage for an analysis of bank profits. Bank profit is identified with homogeneous rent – a misleading approach which becomes even more problematic in the context of the ongoing transformation of banking. This section will argue that some interesting insights into bank profit can be gained from the discussion on financial profit in the literature on financialization. This chapter does not aim at reviewing the vast literature on financialization which comprises a variety of issues, including changes in non-financial

corporations, households, and financial system itself. A review of the broader debate can be found in Krippner (2005), Orhangazi (2007), van Treeck (2009) and Stockhammer (2010). This section will focus only on the question of financial profit as it is discussed in the context of the financialization debate.

Contrary to the mainstream literature, the question of financial profit occupies a central position in the heterodox literature on financialization. It can probably be explained by its closeness to the political economy tradition a major concern of which has always been the source of surplus. As a result, a systematic rise in financial profits has become one of the main themes in the literature on financialization. Krippner was the first to define financialization as “a pattern of accumulation in which profits accrue primarily through financial channels rather than through trade and commodity production” (Krippner 2005, p. 174). Epstein and Jayadev (2005) have also documented a systematic increase in the financial profit in the major OECD countries since the 1980s. More recently, a Marxist approach to financialization as a rise in financial profits was developed by Lapavistas (2009) and dos Santos (2009).

In the financialization literature financial profit is often identified with rentier income, which resembles the treatment of bank profit within the monetary theory of production approach discussed above. This strand of the debate can be traced back to Epstein and Power (2003) and Epstein and Jayadev (2005) who adopt Kalecki’s vision of rentier income as that received by owners of financial firms and returns to financial assets in general (Kalecki 1990 [1943]). Epstein and Jayadev measure rentier income as a sum of profits of the financial system firms and interest income of non-financial private units. They study rentier income in the OECD countries in 1960-90s and conclude that the

share of rentier income in GDP is higher in 1980-90s than in 1960-70s. In general, the take on financial profit as rentier income is typical in the post-Keynesian literature (2010, Stockhammer 2004).

The origin of identifying financial profit with rent by these authors can be probably traced back to Keynes who considered a depression of interest rates to be the main mechanism of “the euthanasia of the rentier, of the functionless investor” (Keynes 1964 [1936], p. 376). Given that the early 1980s – the starting point of the dramatic rise in financial profits – were characterized by a drastic rise in the interest rates, that is, the opposite of Keynes’ dream, it is not surprising that the rise in financial profit came to be identified with a “resurrection of the renter” (Pollin 2007) and the financial profit – with rent.

This approach to financial profit is problematic for a number of reasons. First, it does not distinguish between different types of financial profit indiscriminately identifying it with homogeneous rent. Second, although it correctly notices that financial profit has been accruing to different sectors of the economy, it conflates these forms of profit by turning them all into a homogeneous category. It is misleading, as bank profit is conceptually distinct from interest and dividends accruing to, say, households. Third, after having peaked in 1982, the interest rates have been falling in the US, whereas profits of the financial sector have been rising, both absolutely and as a share of total domestic profits. It implies that it is probably misleading to seek the key to understanding the financial profits in the dynamics of interest rates. With this, a possible link to Keynes’ argument about rentier disappears, and another set of analytical tools becomes necessary for an analysis of financial profit.

In addition to the discussion of financial profit as rentier income, there has been another strand in the financialization literature that can offer other insights relevant for the purposes of the present study. It concerns the macroeconomic sources of the financial profit.

The origin of this literature can be identified with *The Long Twentieth Century* by Arrighi (1994). Following Braudel's *Capitalism and Civilization*, Arrighi looks at financialization not as a particular stage of the capitalist development, even less its final stage as in Hilferding's *Finance Capital*, but rather as a recurrent phenomenon of financial expansions signaling a change in the regime of accumulation of a particular country where it takes place and a change in the hegemonic power in the world economy as a whole. Arrighi (1994, p. 221-238) argues that a phase of material expansion is necessarily followed by intensified competition which in turn, according to Adam Smith, leads to a decline in the profit rate. Given the nature of capitalist agencies that seek profit maximization regardless of a particular source where these profits come from, they compare profit from different types of activities (specifically, Marx's M-C-M' and M-M'). Therefore, the decline in profitability of productive activities causes capital to "switch from trade and production to financial speculation and intermediation" (Arrighi 1994, p. 221)⁵⁰. This leads to a financial expansion with profits coming from "financial deals" (Arrighi 1994, p. 229), or "speculative profits" (Arrighi 1994, p. 235). Arrighi illustrates how this pattern was followed by Genoese, Dutch, British and the US capital

⁵⁰ In spite of his insight into the financial profit being qualitatively different from profit in the sphere of production, throughout his work Arrighi mistakenly conflates the spheres of production and circulation of commodities, hence, profit from production and trade.

resulting in four regimes of capital accumulation with a hegemonic power corresponding to each of them.

Arrighi is right in his emphasis on the qualitative difference between financial transactions and production, therefore, between the two associated forms of profit. Nevertheless, his approach is deeply problematic. First, he uses the terms financial deals, lending, and speculations interchangeably. He does not seem to recognize a qualitative difference between lending and capital investment, on the one hand, and other forms of financial activities, such as trade in securities, for instance, on the other hand. The switches in the regime of accumulation he discusses are often associated with the former, but not the latter. Second, Arrighi's approach can explain a change in activities of a given group of capitalists and related changes in the spatial configuration of the hegemonic powers, but it cannot apply to the world economy as a whole. The fact that a sub-section of the capitalist class comes to "profit from enterprises organized by others" (Arrighi 1994, p. 140, 150) does not say anything about the ultimate source of their profit. Therefore, Arrighi does not explain the macroeconomic source of the rising financial profits associated with the financial expansions. This has become the main object of criticism of Arrighi by Pollin (1996).

An approach related to Arrighi was developed by Krippner (2005). She puts the question of financial profits at the heart of her study by defining financialization as "a pattern of accumulation in which profits accrue primarily through financial channels rather than through trade and commodity production" (Krippner 2005, p. 174, see also p. 181-182).⁵¹ Her "accumulation-centered approach" rightly notices a fundamental change

⁵¹ By financial channels Krippner means activities related to provision of liquid capital in expectation of future dividends, interest, or capital gains.

in the channels of profit extraction and an expansion of financial profits associated with financialization. She was one of the first to document empirically that this expansion is driven by, first, the rising profit share accruing to financial institutions and, second, a changing behavior of non-financial corporations that have come to increasingly rely on portfolio income – dividends, interest payments, and capital gains – as the source of profit. Nevertheless, Krippner does not distinguish between different forms of financial profit. Specifically, she does not discuss the conceptual differences between profit from holding financial assets (dividends and interest) and profit from provision of specific financial services. This is especially important in the light of the other limitation of her approach.

Krippner, similar to Arrighi, does not address the question of the macroeconomic sources of financial profits. Although she argues that her concern is “where profits are generated” (Krippner 2005, p. 175), which creates an impression that her focus is precisely on their macroeconomic source, her actual analysis deals not with the ultimate origin of these profits and not even mechanisms of their extraction, but with documenting their rising share in the US economy. Therefore, after having posed a right question of the sources and origin of profits, Krippner seems to confuse the process of value production and its redistribution. Put differently, when she argues that profits increasingly accrue through financial channels, it is unclear whether it implies that financial transactions increasingly contribute to value creation, i.e. generate this value, or they have become an increasingly important mechanism of redistribution of the already existing value. The difference between different forms of financial profits – portfolio

income and profit from providing financial services – would be decisive here, as these forms of profit might have different underlying macroeconomic sources.

Pollin was the first to challenge the recent discussion about the so-called profits from pure financial deals and profits accruing through financial channels by posing the question of the macroeconomic sources of these profits. His insights constitute one of the main foundations of the present study that will attempt to answer Pollin's question to Arrighi in the context of the transformation of banking: Where do financial profits from the "transformed banking" come from?

Pollin challenges the meaning of Arrighi's "financial expansions" and "pure financial deals" and argues that "Arrighi never explicitly poses the most basic question about the $M \rightarrow M'$ circuit, which is, where do the profits come from if not from the production and exchange of commodities?" (Pollin 1996, p. 115). In his own response to this question, Pollin singles out three possible sources of sustainable financial revenues. The first one is redistribution within the capitalist class – a zero-sum game resulting in no profit from financial transactions for the economy as a whole. The second case takes place when financial transactions are accompanied by a change in the balance of power in favor of the capitalist class as a whole, with corporate takeovers followed by forcing down wages and taxes being an example of that. This results in redistribution of income in favor of the capitalist class as a whole. Finally, financial mechanisms can allow for channeling funds into more profitable investment opportunities, thus, can help raise the surplus produced in the sphere of material production and exchange⁵². Put differently, in this case the $M \rightarrow M'$ circuit operates successfully only because there is an underlying M

⁵² Pollin argues that this is the actual object of Arrighi's analysis.

→ C → M' circuit. Pollin concludes that one needs carefully to distinguish between these three distinct sources of financial profit – profit from M → M' transactions.

In his response to Pollin, Arrighi defines the periods of financial expansion as those “in which profits come, *not* from the further expansion of trade and production, but from borrowing, lending and speculating (as encapsulated in Marx’s abridged formula of capital M→ M’)” (Arrighi 1997, p. 154). He insists that

not only does the account of historical capitalism proposed in *The Long Twentieth Century* identify the three different sources of financial profit that Pollin claims I failed to distinguish, in addition, it assigns to each source a distinct role in the dynamics of financial expansions. (Arrighi 1997, p. 157)

Specifically, Arrighi identifies the first type of profit in Pollin’s taxonomy with the cut-throat competition characteristic of the period of capital over-accumulation. Capitalists come to invade each other’s sphere of operation, so that losses of one become a condition of profit for others, resulting is the zero-sum game among them. These developments create the supply conditions for a financial expansion – tendency to keep profits in a liquid form, instead of reinvesting them in the process of production. These supply conditions are then matched by the demand conditions operating on the site of inter-state competition for liquid funds, with the two coupled together generating the financial expansion and a specific type of profit associated with it. This profit can be thought of as the second form mentioned by Pollin and it comes from redistribution “from all kinds of communities to capitalist agencies” (Arrighi 1997, p. 156). For Arrighi, these communities include both the state and working-class communities. Finally, he argues that the third type of profit named by Pollin – that from shifting uses of funds to more profitable investment opportunities – emerges only at the stage of supersession of the financial expansion by another stage of material production. Thus,

Arrighi claims that material production is not an underlying source of the financial revenue at the stage of financial expansion, but rather a consequence of an end of this expansion.⁵³

The Pollin-Arrighi debate is relevant for the purposes of the present study. Pollin is right in explicitly posing the question of the macroeconomic sources of financial profits. His linking all forms of financial profit to flows of value in one way or another is also paramount. At the same time, Arrighi's emphasis that during financial expansions profits primarily come from redistribution of the already created income needs to be retained. Finally, in spite of these insights, both Pollin and Arrighi while trying to establish the ultimate sources of profit overlook mediations and intermediate steps that are important in the case of some types of financial profit. The relevance of these mediations will become apparent in chapters 4 and 5.

A line of reasoning similar to Pollin's was advanced by van Treeck (2009) with respect to Krippner. He is critical of her definition of financialization on the grounds that

in a descriptive sense, it is undoubtedly true that many profits are nowadays linked to financial activities. Yet, given the macroeconomic definition of profits... it may be helpful to recall that firms in the aggregate can by no means autonomously choose either between real investment (production) and profits at large or even between non-financial and financial profits. (van Treeck 2009, p. 911)

On these grounds van Treeck concludes that it is "at least semantically, if not conceptually, problematic to consider 'the financial sector as a source of profits for the economy'" (van Treeck 2009, p. 911). Thus, he emphasizes that the financial sector cannot generate any profits without the process of production and exchange. For him, it is

⁵³ Here Arrighi fails to see that material production can be a cause of both – sustained profitability of financial transactions and a shift to a new phase of material production.

more adequate to talk about the financial system becoming more successful at extracting profits from the real economy, rather than being their source.

The emphasis by Pollin and van Treeck on the macroeconomic sources of financial profits is important, because it fills a gap in the discussion on bank profit within the monetary theory of production. They are right criticizing Arrighi and Krippner, respectively, for not having addressed the question of the sources of profit from the purely financial transactions. For Pollin and van Treeck, all the profits originate in the sphere of production and exchange, and financial transactions merely redistribute the currently produced national income. This approach is obviously correct when it comes to interest and dividend payments, and these are the only forms of revenue van Treeck has on his mind. Nevertheless, financial transactions, including financial services provided by banks and other financial institutions, include a vast variety of non-lending activities. To name a few, there are trading gains, securitization revenues, fees from mergers and acquisitions, and asset management fees. What is the macroeconomic source of these revenues?

Van Treeck does not address this question, as he does not seem to be concerned with the macroeconomic sources of bank profits. Although he does not define what he means by the financial profits, it appears that for him these are revenues associated with holding financial assets⁵⁴ – dividends and interest payments – that can accrue to all sectors of the economy. Pollin has a broader concern, as he does not narrow his focus to profits from holding financial assets. It might appear that for him the source of profits

⁵⁴ That is why it is not surprising that for van Treeck financial profits can only come from the process of production, as this statement is reduced to a simple accounting identity. Dividends are a part of the output already created, and capital gains are netted out as they do not add value.

from all types of financial transactions, and not only lending and investment, is the current output, unless it is a zero sum game. This is illustrated by his example of profits from corporate takeovers that, in Pollin's view, have often ultimately come from wages via redistribution of income in favor of the capitalist class as a whole (Pollin 1996, p. 115, Arrighi 1997). Thus, Pollin wants to retain the classical political economy conclusion that value is created only in the sphere of production, and the sphere of circulation, including its financial dimension, only redistributes the already created value.

From the perspective of the present study, Pollin and van Treeck are right in posing the question of the macroeconomic sources of financial profits and linking them to flows of value. Nevertheless, they are too quick to dismiss both the idea of the purely financial deals and other potential sources of financial profit. Although one can think of reasons why they are dismissive of these two points.

First, the concept of purely financial deals does not have a definite content in Arrighi who uses it interchangeably with lending and speculation. But lending is not speculation. At the same time, there is a vast array of financial activities that are indeed at one remove from production – trading securities, facilitating mergers and acquisitions, and so on. They might indeed be called “purely financial deals”, and Arrighi's insight about the distance between the process of value creation and financial transactions needs to be retained. By the same token, although Pollin is right that the macroeconomic sources of profits from these activities require an analysis, it is not enough to postulate that they come from either wages or profits. Even if it is the case, it remains to be shown.

Second, it is not surprising that for Pollin and van Treeck only the current output, i.e. profits and wages, can be sources of financial profits, given that in both post-

Keynesian and neoclassical theories there is no other macroeconomic aggregate that can be an independent source of profit. Van Treeck (2009, p. 908) is very clear about it arguing that “macroeconomic profits must always be based on real income flows”. Therefore, Pollin and van Treeck face an objective limitation in the form of a lacking analytical framework and tools for an analysis of financial profit as distinct from the current flows of value. It will be argued in chapter 4 that a Marxist theory has a category of loanable capital that can be such a macroeconomic aggregate and which can constitute another independent source of profit. It would give an analytical advantage for an analysis of financial profits to a Marxist theoretical framework.

To recap, the discussion on financial profits in the literature on financialization can offer important insights for understanding bank profit. Although in this literature financial profit is often identified with the rentier income, similar to bank profit within the circuitist approach, there are other takes on financial profit. An important angle was suggested by Pollin and later by van Treeck, who emphasize the macroeconomic sources of financial profit, although they do not address bank profit in particular. The present study will attempt to bridge the gap between the two streams of the literature – on the bank profit and on the financial profit in general. In spite of offering important insights, Pollin and van Treeck are too quick to dismiss Arrighi’s idea of the purely financial deals, although for a good reason. First, Arrighi is not clear what he means by it, which does not deny that there are financial transactions that are indeed at one remove from the sphere of production, which would have implications for the nature and sources of profit from these deals. Second, Pollin and van Treeck face objective limitations of their

analytical apparatus, in which revenues – unless they are a zero-sum game – can only come from redistribution of wages or profits.

3.6. Conclusion

It was argued in the present chapter that the approach to banking within the monetary theory of production offers some insights that can overcome the limitations of the mainstream theories of banking and bank profit. Contrary to the old mainstream approaches, the circuitist take on banking places banks on a foundation other than information asymmetries and market imperfections that have been challenged by the ongoing transformation of banking. Furthermore, unlike both the old and the new mainstream theories, the monetary theory of production acknowledges difference across economic sectors and explicitly posits banks in their relationship to these sectors.

Thus, at a first glance the take on banking within the monetary theory of production seems to be well suited for providing a theoretical foundation for understanding banking and bank profits. Nevertheless, a closer look at this theory uncovers its limitations and shows that it is a partial theory of banking that, similar to the mainstream theories, encounters its limits while trying to explain the ongoing transformation of banking. By the same token, the take on bank profits within the monetary theory of production was shown to be inadequate for the purposes of the present study. To address the problematic treatment of bank profits, some insights can be gained from a broader discussion of financial profit in the literature on financialization. The emphasis on the macroeconomic sources of financial profits is shown to be paramount.

On the basis of the discussion of the theories of banking and bank profit in chapters 2 and 3, it has become clear that in order to address the transformation of banking and bank profits, a theory of banking should satisfy a series of requirements. These can be summarized as follows.

First, similarly to the circuitist take on banking, a theory of banking should posit banks in their relationship to the circuit of capital and revenue. Nevertheless, contrary to the monetary theory of production, the essence of banking business in general cannot be defined through a particular link to these circuits.

Second, banks should not be theorized as essentially phenomena of information-collection, risk management, or development. Banks are primarily a monetary phenomenon, and Keynes' emphasis on liquidity provision can offer a take on a trans-historical function of banking. Money creation *ex nihilo* is merely a particular form of this function of banking. Nevertheless, liquidity provision *per se* also does not offer a sufficient foundation for theorizing banking, and an emphasis should rather be shifted on the counterpart of the act of liquidity provision, namely, extension of a promise to pay. It would give more mileage for understanding social relations involved in financial transactions.

These two aspects highlight the role of the method of abstraction. Banks should be theorized not at the level of abstraction of money, but rather of a particular form of capital with its own circuit. By the same token, this abstraction should pivot on the simplest form of the banking business that contains a possibility to be developed into its more concrete forms.

Third, bank profits are paramount. It was argued that they can be a good lens for showing that banks have not been dying out and for studying the content of their transformation. Nevertheless, the mainstream literature has a strong empiricist bent. As a result, it merely lists a number of types of bank profits as they are given by the accounting standards, without questioning their content and the underlying social relations. The multiplicity of forms of bank profits is the only insight that can be gained from this literature. By the same token, the circuitist theory of banking also fails to offer a satisfactory theory of bank profits as they are analyzed as a homogeneous rent. Therefore, a theory of banking should allow for a multiplicity of forms of bank profits, at the same time it should pose them on a theoretical foundation and show their macroeconomic sources and the underlying social relations, as stressed in the literature on financialization.

Chapter 4 will discuss a possible theory of banking and bank profits that would satisfy these requirements.

CHAPTER 4

MARXIST THEORY OF BANKING AND BANK PROFIT

4.1. Introduction

It was shown in the previous chapters that banks have been changing over the past few decades. These changes have been documented in the vast empirical literature. Moreover, as discussed in chapter 2, theoretical literature has also shown signs of awareness of these trends. In the course of this transformation, the mainstream theory of banking has encountered its limits. Therefore, there have emerged attempts to put banking on a foundation other than information asymmetries and market imperfections. These attempts addressed various aspects of the transformation, but no coherent mainstream theoretical framework has emerged to replace the information-theoretic approach. Chapter 3 argued that the heterodox theories have never theorized banking through information asymmetries, and they also explicitly locate banking among the other sectors of the economy. Nevertheless, the exact way they do it is self-defeating, and the heterodox theories fail to provide a foundation for understanding banking and its transformation.

On this basis chapter 3 concluded with requirements to a theory of banking that would provide a broad analytical framework for understanding banking, its transformation, and bank profits. Given that the mainstream and heterodox theories do not offer such a framework, does it mean there is a need for a new theory of banking and bank profit?

It will be argued in the present chapter that a Marxist theory can offer such a theoretical framework. Like the new mainstream theorizations and heterodox approaches

to banking, it offers a broad theoretical foundation for understanding the content of banking activities. This content is different from information asymmetries and market imperfections that have been challenged by the transformation of banking. Yet, contrary to both the old and the new mainstream theories, it explicitly locates banking among the sectors of the economy, but does it in a way allowing for multiple links to the circuit. Thus, unlike the heterodox approaches discussed before, it is better suited to explain the transformation and it does so retaining the strengths of the discussion on financial profit in the literature on financialization. As a result, a Marxist theory of banking can nest other approaches to banking and their major insights, but also overcome their limitations. Therefore, it also provides a better basis for understanding the nature of bank profits and changes in them.

The chapter is organized as follows. Section 2 lays out Marx's theory of banking by locating it in the approaches to banking that have existed in the history of economic thought – the bills view and the goldsmiths view. Marx is shown to have adopted the general theoretical framework of the bills view, yet to have fundamentally altered it to explicitly posit banks in their relationship to the circuit of capital and revenues. Section 3 shows the mileage this theory gives for understanding the ongoing transformation of banking. It is argued that contrary to the widely spread belief, some of the changes in banking are not unprecedented. At the same time, other changes – that have attracted less attention, especially in the mainstream theories – are more profound, and a Marxist theory gives a basis for understanding them. Finally, it is argued what seems to be a small change within the bills view can imply a more profound change that can be recognized due to Marx's emphasis on the link between banks and the circuit of capital. Section 4

compares a Marxist theory of banking and transformation of banking with other theories and shows that the former is a more general theory and how exactly it incorporates the major insights of other approaches. Section 5 lays out a general theoretical framework for understanding bank profits. This framework is grounded in the Marx-Hilferding theory of bank profit, but it advances this theory by, first, further developing the concept of loanable capital and, second, distinguishing between two key types of financial profit – profit from holding promises to pay and profit from handling (dealing in) them. It is argued that the latter is a special type of financial profit that is conceptually distinct from both profit in the sphere of production and interest. The section lays out a general theory of this form of profit and prepares the ground for a discussion of concrete forms taken by this type of profit to be discussed in the following chapter. Section 6 concludes.

4.2. Bills View vs. Goldsmiths View

There are two approaches to banking in the history of economic thought – what can be called the bills view and the goldsmiths' view. According to the bills view, banks emerge from discounting instruments of commercial credit, mostly, bills of exchange. This serves as a basis for their issuing bank notes in exchange for the bills they discount. By doing it, banks substitute one promise to pay by another which is more widely accepted as a means of circulation. As a result, the banking business of discounting generalizes the credit relations beyond the counterparties exchanging the commodities. Within the bills view, this activity of issuing one promise to pay in exchange for another

is viewed as the essence of banking business, with all the other banking activities developing out of this simplest form.⁵⁵

Adam Smith can be considered an early proponent of the bills view. He first mentions banks in the context of paper money, arguing that promissory notes issued by banks are the most well known and, in fact, the most suitable form of paper money substituting for gold and silver money (Smith 1884, p. 118). This way of introducing banks might be interpreted as his staying at the origin of theories of banking based on money creation *ex nihilo*. Nevertheless, for the purposes of the present study it is decisive that although Smith does indeed introduce banks in the context of money creation, he does it not by retreating to an abstract process of money creation *ex nihilo*. Contrary to the theories discussed in the previous chapter, Smith connects banks issuance of promissory notes with discounting (Smith 1884, p. 121). His discussion of the actual historical origin of banks pivots on the establishment of Banks of Deposit by the city of Amsterdam and other small states⁵⁶ in the context of this state policy aiming to protect the interests of domestic merchants who would otherwise face unfavorable terms of discounting due to uncertain value of the national currency (Smith 1884, p. 194-198). Thus, the trust that emerges with establishing large banks aims to substitute for a lack of

⁵⁵ It is worth noting that Schumpeter's treatment of the two approaches to banking is confusing. Only once he rightly refers to the commercial-bills view as the opposite of the goldsmiths' view (Schumpeter 1954, p. 317). Elsewhere, he conflates the former with banking as intermediation (Schumpeter 1954, p. 729-730, 1111), and on these grounds is highly critical of it, which is not surprising given that he was a strong advocate of the emphasis on money *ex nihilo* (Schumpeter 1961 [1911/1934]). In our view, Schumpeter misunderstood the gist of the commercial-bills view and overlooked that its core is not the intermediary function of banking, but rather the business of exchanging promises to pay as the simplest form of the banking business. This exchange does not hinge on either intermediation function of banking, or its capacity to create money *ex nihilo*. Put differently, bills can be discounted by either cash deposited with banks or bank notes, with the two being concrete forms of doing what essentially constitutes the banking business.

⁵⁶ Among them were also Venice, Genoa, Hamburg, and Nuremberg.

trust in national currencies of the small states of origin of these banks. Taken in this context, the link to discounting operations and exchange of promises to pay is again apparent.⁵⁷

The emphasis on banks' discounting operations gives Smith a strong starting point for a study of banking, because, first, it immediately puts the question of money creation in the context of a swap of two promises to pay – merchants' bills and bank notes. It has nothing to do with an abstract activity of money creation, which obtains its content from the needs of commerce, trade, and later industrial capital. Therefore, it allows one to look at banking business, even in its simplest form of discounting, as an exchange of promises to pay. Second, Smith treats promissory notes as a way for banks to increase the scope of their operations, hence, raise their profits. Through issuing notes, a banker is "enabled to make his clear gain of interest on so much a larger sum" (Smith 1884, p. 121). This in turn brings to the surface the active principle governing the banking business. Banks are capitalist firms aiming at increasing their profits⁵⁸, and issuing own promissory notes to discount bills of exchange and granting cash accounts become but one way of achieving this goal.

⁵⁷ Considering the historical origin of banking, Smith also has a lengthy discussion of banks as deposit takers. It might be interpreted as his advocating for treating banks as intermediaries. Therefore, one might argue that there is a clash between Smith's treatment of the logical and the historical origin of banking. In his discussion of the former, he might be interpreted as being concerned with banks as money creators, in the latter – with their intermediation function. Nevertheless, from the perspective of the present study, there is no clash between the two approaches to banking in Smith. The unifying principle behind these two parts of Smith's theory is in his identifying banking with an exchange of promises to pay that explains both banks' strive to create money and the small states establishing Banks of Deposit.

⁵⁸ As Smith put it, "public utility, however, and not revenue, was the original object of this institution [bank]. Its object was to relieve the merchants from the inconvenience of a disadvantageous exchange. The revenue which has arisen from it was unforeseen, and may be considered as accidental" (Smith 1884, p. 198). The profit motive of banks is apparent, although, contrary to Smith, there is obviously nothing accidental about banks becoming profit-seeking enterprises.

This general framework of tracing the origin of bank credit to the spontaneously arising commercial credit, and the origin of bank notes – to commercial promises to pay, pivoting on the active principle behind the banking business, was also adopted by Thornton (1802, p. 37). In a study of the origin of the money market, Bagehot has also stressed that banks could not have even emerged unless they were active issuers of notes, contrary to passive collectors of deposits. For him, “in the issue of notes the banker, the person to be most benefited, can do something... But in the getting of deposits he is passive” (Bagehot 1877, p. 86). According to Schumpeter, bills view was also adopted by the Banking School (Schumpeter 1954, p. 729-730). Willis (1921, p. 3) has summarised this view by defining the function of banking as “guaranteeing the limited or individual purchasing power represented by the obligation of each individual, by accepting it and substituting in lieu thereof the bank’s own obligation”.

The second approach to banking is represented by what is called the goldsmith’s view (Schumpeter 1954, p. 317), or deposit banking (Bagehot 1877, p. 75), according to which banks have emerged as safe keepers of hoards, or custodians. They have issued notes to certify their acceptance of deposits, which is the embryo form of bank notes. Accumulation of large amounts of deposits has made banks seek profitable outlets, which is the origin of their lending business. From this perspective, the essence of banking business consists in collecting deposits and making loans, that is, in passive intermediation. Cantillon was an early advocate of this view. For him, banks collect deposits and make loans after having secured the necessary reserves (Cantillon 2001 [1755/1931], p. 120-121). This allows them to accelerate the circulation of money. According to Schumpeter (1954, p. 319), similar ideas were expressed by Galiani and

Turgot. Rist is viewed by Schumpeter as the leading exponent of this view in the 20th century. It has also become the dominant approach within the mainstream theories of banking.

Marx has adopted the general theoretical framework of the bills view, but he has also fundamentally altered it. As it is always the case with the bills view, for Marx, commercial credit which spontaneously emerges among industrial capitalists forms the foundation of the banking credit (Marx 1977 [1894], p. 479). As elaborated by Itoh and Lapavistas (1999, chap. 4), the latter emerges as a way to overcome the limits of the former. Yet Marx's fundamental contribution to the theory of banking is his emphasis on the difference between credit relations in general and the specific form they take within the capitalist mode of production. He discusses it within the context of the difference between usury and interest-bearing capital. As capitalism develops, it poses the credit relationship on the capitalist foundation and subsumes it to the purposes of value creation. Thus, the difference between usury and interest-bearing capital amounts to the difference not in the act of lending per se, but rather in the conditions under which lending takes place and characteristics of the lenders and the borrowers. As he put it,

What distinguishes interest-bearing capital – in so far as it is an essential element of the capitalist mode of production – from usurer's capital is by no means the nature or character of this capital itself. It is merely the altered conditions under which it operates, and consequently also the totally transformed character of the borrower who confronts the money-lender (Marx 1977 [1894], p. 600).

Thus Marx preserves the main pillar of the bills view by treating the essence of the banking business as an exchange of promises to pay. Hence, he also retains the emphasis on the inherent flexibility of this business, as promises to pay can take multiple forms, can be exchanged in various ways and on different conditions. Nevertheless, in

line with his general contribution to political economy, Marx moves away from an ahistorical theory of banking characteristic of the bills view as it existed at his time. Instead, he explicitly locates banks within the capitalist mode of production.⁵⁹

In spite of its origin more than a century ago, the Marx's version of the bills view can be a powerful foundation for understanding modern banking. It involves two aspects. First, the multiplicity of forms of financial transactions is inherent to the bills view, yet there is a unifying principle behind these many forms. An exchange of promises to pay lying underneath the act of discounting remains the content of more complex financial transactions. It was emphasized by Hicks, for whom

financial transactions are always, in some sense or other, loans; so the simplest form of loan contract – money being paid over now, in return for a promise of repayment, with interest, at some future date or dates – is the element from which we should start. One could start that way, and go on to admit that the amounts, and dates, of repayment may be not fixed but conditional on things that may happen in the future; so proceeding to insurance contracts, subscription to equities, and so on (see also Hicks 1967, p. 53, Hicks 1989, p. 50).

According to this approach, the content of all financial transactions involves an exchange of money for a promise to pay, so that a simple credit relation can be viewed as an embryo form of more complex forms, both debt- and equity-like. Then banking business can be understood as the business of exchanging one set of promises to pay for another, in different ways, on different conditions and terms, and using different instruments. This approach is different from Itoh and Lapavistas (1999, p. 82) who also look at discounting as the simplest form of credit relations and define the content of banking as an exchange of promises to pay. Nevertheless, for them the difference

⁵⁹ Incidentally, the emphasis on locating banks within the mode of production is what distinguishes Marx from more recent versions of the bills view, such as, for instance, developed by the Austrians and Hicks. Both Marx and the Austrians stress the spontaneous origin of the credit relations, but the former brings in the role of the mode of production thus radically differentiating himself from the latter.

between debt and equity is decisive and needs to be retained. Although Hicks is right that a credit relationship is a fruitful starting point for an analysis of more complex financial deals, there are significant differences between loans and, say, equities. The latter is a title of ownership that does not involve a promise to pay back the principal amount invested. Furthermore, from a broader perspective, a difference between certain financial instruments can be important as it can reflect differences in social relations, changing structural or institutional characteristics of an economy. For example, Hilferding (1910, p. 107-109, 121-123) stressed that joint-stock capital corresponds to a new set of social relations reflecting changing conditions in capital accumulation, namely, separation of management and ownership associated with a rising scale and fixed capital investment requirements. For the purposes of the present study, the question of joint-stock capital in its relationship to capital accumulation is of secondary importance. Therefore, we will adopt the general framework of the bills view broadly defined and treat all financial transactions as an exchange of money for a promise to pay.

Second, the emphasis on the active principle endemic to banking since its origin can shed the light on the ongoing transformation of banking and put it in a broader historical context. This deserves a closer analysis to be undertaken in the following section.

4.3. Transformation of Banking

Marx's approach to banking opens a room for two types of transformation, namely, transformation of the specific ways in which promises are exchanged, i.e. in the mode of performing banking business, and, second, in the way banks relate to the mode

of production and the circuit of capital. The general framework of the bills view provides a basis for understanding the former, while Marx's contribution to it by locating banks in the mode of production sheds the light on the latter. The strength of a Marxist approach is in its ability to capture both. As was shown above, while the first transformation is often discussed in the mainstream theories, the second one is usually overlooked. The latter is recognized only by the heterodox theories of banking and, more broadly, in the literature on financialization, yet its theorization is misleading due to the self-defeating way the connection of banks to the circuit is specified. Consider the two types of transformation in a greater detail.

Transformation I. The transformation of the ways in which banking business is performed has a long history. In a sense, the history of banking is the history of the multiplicity of types of instruments banks deal in, and specific modes and conditions of handling these instruments. It obviously deserves a separate historical and empirical examination, which is beyond the scope of the present study. What is important for our purposes is that behind the multiplicity of forms the Marx's theory of banking recognizes a general logic of development of the banking business. Given that banks are independent capitalist enterprises their functioning is driven by the profit motive.⁶⁰ Their profit is positively related to the volume of their activity, hence, there is a tendency for banks to strive to increase their lending capacity. If banks were to lend only their own capital, the amount of lending would be limited. The traditional ways to increase the lending capacity

⁶⁰ The active principle behind banking has always been emphasized by Minsky (1993, p. 14) for whom banks are profit-seeking firms, "they are not charitable institutions".

is to issue bank notes and to attract deposits (Itoh, Lapavitsas 1999, p. 94-95, Lapavitsas 2003, p. 79, Lapavitsas 2007, p. 426)⁶¹.

Thus, banks come to create money and become intermediaries. But contrary to the circuitist and the mainstream approaches, respectively, neither of these two functions can be viewed as the essence of banking, hence, neither can be a starting point of analysis of banking. Rather, deposit taking and money creation develop out of banks being capitalist enterprises striving for profit which in turn depends on the scale of their lending operations. These two activities are essentially two original forms of financial innovation, and they do not appear innovative only from the perspective of the 21st century. An important consideration here is that, once developed, there is neither logical nor historical reason to assume that one of these functions would disappear. Even when banks develop mechanisms of further raising their lending capacity, money creation ex nihilo and deposit taking remain. A casual observation confirms that even modern banks engaging in liability management, asset sales, and securities trading, continue to hold demand deposits and issue bank drafts. Hence, the description of the ongoing transformation as disintermediation is not quite accurate. Banks remain intermediaries, but also develop additional modes of performing banking business.

Once deposit taking and money creation are theorized merely as particular ways in which banks increase their lending capacity, the ongoing transformation of banking does not appear any more as an unprecedented change. Rather, it is a concrete form of the same strive to increase the lending capacity that has existed since the origin of banking, although under altered historical conditions, on a different scale and using more advanced

⁶¹ An increase in the lending capacity as a major concern of banks is also emphasized by Cutler, Hindess, Hirst, and Hussain (1977, p. 102).

methods and technologies. This holds for the major changes in banking, both on the asset and liability sides.

First, liability management that emerged in the US in the early 1960s with an invention of marketable certificates of deposits (CD) is often treated as a profound change in banking business (e.g., Moore 1988, p. 27). Later on, borrowed funds have taken the form of the Eurodollars, commercial paper, federal funds, repurchase agreements (repo), and other non-deposit liabilities, whose share in bank liabilities has been growing since then. In spite of the multiplicity and complexity of these instruments, it is misleading to treat this form of liability management as a fundamental shift in banking, as in their essence attracting deposits or money creation constitute two original ways of managing bank liabilities. In all the cases, banks seek liabilities to match their assets. By the same token, in all the cases the bank activity is driven by a motive to increase lending capacity or replenish reserves. In spite of these similarities, there is a difference in the source of liabilities, as deposit taking represents collecting idle funds scattered across the society, issuing own liabilities does not hinge on anyone's assets and is based on demand deposits becoming money-like (Hicks 1989, p. 55-58), while wholesale funding rests on attracting liquidity from the market. The latter obviously depends on the existence of the developed money and capital markets, hence, cannot emerge at an early stage of financial development.

Second, loan sales and securitization represent the same logic of raising lending capacity. They differ from liability management in that they operate through the asset, not liability, side of the bank balance sheets, nevertheless, they reflect the same principle of banks seeking ways to increase their lending potential. As a result of operating through

the asset side, the other major difference between the liability management and the loan sales is that the latter frees up banks from the other major concern of the banking business – adjusting reserves to match up to the increased liabilities. Therefore, it is superior to liability management, from the banks’ perspective. The fact that the special purpose vehicles (SPV) have not been required to hold reserves against their liabilities, in the case of the loan sales the question of costly reserves is removed for the financial system as a whole, not only for the banks originating a loan. Incidentally, another widely spread argument about the novelty of securitization stresses a radical transformation from loans being non-marketable assets to their becoming a security, like any other. It is not quite correct, as, first, historically banks have emerged as holders of marketable assets, bills of exchange, and in that sense there is nothing extraordinary in marketability of bank assets in principle. In this sense marketability of loans should be treated as a challenge not to the banking business in general, but rather to a particular form of banking business – that based on the relationship lending and commitment which used to be dominant in the bank-based financial systems. Second, the public confidence in banking – that periodically gets shaken by crises – is the basis for marketability of loans. In that sense the socialization and objectification of the credit relationship that emerges with the development of the pyramid-like financial system stressed by Itoh and Lapavitsas (1999, p. 98) and Lapavitsas (2003, p. 70-71), becomes a foundation for the possibility of loan sales. Put differently, if bills of exchange held by banks were originally marketable on the basis of the productive activity of the capitalists involved, bank loans acquire the same marketability through the functioning of the financial system itself. In this sense, the social foundations of acceptability of bank liabilities as means of circulation and of

marketability of their assets are not drastically different from each other, and both are actively shaped by the financial system itself.

To recap, banks are self-standing capitalist enterprises governed by the profit motive. Since their origin, they have been seeking ways to increase their lending capacity, subject to the need to hold reserves. Deposit taking and issuing bank notes were two original forms of banks accomplishing this goal. Once developed, these functions of banking remain, even in developed financial systems. Nevertheless, there is no reason to assume that these would be the only forms of banks' increasing their lending capacity. In general, specific mechanisms of achieving this goal would vary depending on the concrete historical, institutional, and other country-specific conditions. Nevertheless, as a general rule, as the financial system expands, the motive to further raise lending capacity does not go away, whilst the demand for wholesale funds on a short notice increases. Thus, more developed forms of banks' increasing their lending capacity emerge, such as, for instance, liability management and loan sales and securitization. Contrary to the wide spread belief, the proliferation of these activities per se⁶² should not be considered as a radical transformation of banking, but rather as a new historical form of the same logic of banking aiming at raising lending capacity to increase profit.

Transformation II. Marx put the theory of banking on a dual foundation by adopting the general framework of the bills view yet locating banks within the capitalist mode of production. As a result, in addition to changes in the concrete mechanisms,

⁶² It will be argued below that although it is misleading to view liability management and loan sales per se as a fundamental transformation of banking, contrary to the widely spread approach, these changes in banking involve some other transformations that do represent a radical change in banking, specifically, in the sources of bank profit. This latter change is much more fundamental than the former, yet it is overlooked in the literature.

instruments, and related conditions of performing the business of banking, there emerges a possibility for another type of the transformation of banking. It concerns a change in the prevailing link between banks and the circuit of capital and revenue.

The traditional banking activity in the capitalist economies used to be lending to non-financial enterprises. Lending to industrial and commercial capitalists signifies that the credit system was placed to the needs of the capitalist mode of production, and credit was mainly used for the creation and realization of value. Incidentally, this is in line with Marx's derivation of interest-bearing capital through lending in order to embark on the circuit of capital (Marx 1977 [1894], chap. 21, p. 340, 349).

Nevertheless, given that banks are self-standing capitalist firms operating in pursuit of profit, there is no reason to assume that productive lending would be the only type of lending banks would engage in. That is why Marx's emphasis on the characteristics of the borrower discussed above becomes important (Marx 1977 [1894], p. 600). The act of borrowing can remain the same, but it can have a different content depending on the social relations underlying this borrowing and characteristics of the borrower. Incidentally, the fact that the act of lending remains the same regardless of the nature of the borrower explains why this type of transformation remains unrecognized by the mainstream theories that do not capture the social dimension attached to banking activities.

The concept of financial expropriation or exploitation, advanced by Lapavistas (2009) and dos Santos (2009) should be placed in this context. Following Marx, they argue that lending to households is qualitatively different from lending to industrial and commercial capitalists for a number of reasons. First, the former approach borrowing

from the perspective of use-values, whereas the latter from the viewpoint of value (Lapavitsas 2009, p. 132). That is, the former borrow to secure access to basic necessities, such as housing, education, healthcare, and consumer goods, whereas the latter borrow to embark on a circuit of capital and extract surplus value. As a result, second, lending to industrial and commercial capitalists helps create the source of its own repayment, whereas lending to households normally does not (Dos Santos 2009, p. 191).⁶³ This argument is in line with Adam Smith's distinction between lending for production and consumption. For him, borrowing for production allows the borrower to "both restore the capital, and pay the interest, without alienating or encroaching upon any other source of revenue" (Smith 1884, p. 144). By contrast, a borrower for consumption can "neither restore the capital nor pay the interest, without either alienating or encroaching upon some other source of revenue" (Smith 1884, p. 144). Therein lies the fundamental difference between the two types of borrowing. For Smith, the latter is the case "where gross usury is out of the question, contrary to the interest of both parties" characteristic of the former (Smith 1884, p. 144).

For Lapavitsas and dos Santos, the shift from lending to firms to lending to households represents a transformation of banking signifying a reinstatement of essentially usurious relations within the capitalist mode of production. This transformation has become possible due to substantial changes in the major capitalist economies associated with a retreat of the public provision of basic necessities

⁶³ Dos Santos considers residential real-estate bubbles as an obvious, yet partial exception, given that they can generate a temporary source of capital gains for households. In addition, another exception when lending to households can help generate the source of its repayment would be borrowing for education, or student loans. But even in this case the relationship between borrowing and repayment of a student loan remains mediated by, among the other things, becoming employed.

(education, healthcare, pensions, and housing).⁶⁴ Repression of the wage share has also played a role, but to a lesser extent, given that the major bulk of borrowing by households in the developed countries has been in the form of mortgages, and not consumer credit.⁶⁵

It is crucial to recognize that this transformation of banking constitutes more than a plain reinstatement of usury. For Marx, there are two aspects of interest-bearing capital that make it a specifically capitalist relationship. First, borrowing to embark on the circuit of industrial and commercial capital is a specifically capitalist type of borrowing (Marx 1977 [1894], p. 349, 352). Second, equally important, although often less appreciated, the circuit of capital systematically generates temporarily idle hoards, leakages of value, that are assembled by the credit system and transformed into loanable capital (Marx 1977 [1894], p. 403, Marx 1967 [1885], p. 78, 83-85, 321-322, 451).⁶⁶ This generation of the idle funds is the other aspect giving the credit system a specifically capitalist character. It is further reinforced by the credit system coming to assemble idle funds of not only capitalists, but also workers, that is, coming to use both the circuit of capital and the circuit of revenue as the basis of its functioning. The shift of lending away from firms to

⁶⁴ As stressed by Lapavitsas (2004, p. 176-177), the role of the State in designing financial systems and institutions should not be underestimated. The modern US financial system is a good example, where the lack of the public provision of the basic human necessities is coupled with implicit and explicit State guarantees for a specific type of banks created for mortgage lending to households – government-sponsored enterprises (GSEs). Thus, the State not only has helped create conditions under which lending to households on a massive scale has become possible, but also in the best mercantilist tradition described by Stuart (1770b, p. 414) has helped establish financial institutions engaged in this type of lending by assuring their solvency. GSEs have become a modern version of development banks – what Stuart (1770b, p. 359) called ‘banks of circulation upon mortgage’ – in the context of financialization. Instead of land, residential houses have become Stuart’s solid property to be ‘melted down’.

⁶⁵ This is not to deny the significance of second mortgages that became a major source of maintaining or increasing consumer spending in the face of wage repression in the 2000s, along with consumer credit in the form of credit card borrowing.

⁶⁶ This point was further elaborated by Hilferding (1910, p. 67-81), Itoh (1988, p. 259-260, 401), Itoh and Lapavitsas (1999, p. 65-69), and more formally by Lapavitsas (2000).

households obviously undermines the first specifically capitalist aspect of the credit system. Nevertheless, the second capitalist foundation of the credit system does not go away. It continues to assemble hoards scattered across the society and use them for profit extraction. Therefore, lending to households on the basis of the capitalist financial system is even more predatory than usury proper. A usurer normally lends out only own capital, whereas modern banks use temporarily idle money capital of the society as a whole. As a result, banks can perform the essentially usurious activity on a larger scale, hence, generate a greater mass of profit than a usurer.

To recap, a Marxist theory of banking can offer a foundation for understanding the content of banking business and its transformation. It allows one to theorize two types of transformation. The one is related to the concrete ways in which the banking activities are performed, and in many cases it is associated with an increase in the lending capacity. Acknowledging the active principle behind banking gives the unifying principle to the multiple banking activities and shows that there is no fundamental change often attributed to the modern banks. The other type of transformation refers to the type of the borrower and is directly related to broader social transformations.

These two types of transformation of banking appear unrelated. This impression is reinforced by the existing literature. For instance, as was shown in chapter 2 above, the mainstream theories focus only on the first type of transformation, with the second one remaining unnoticed. Chapter 3 argued that heterodox theories of banking recognize both changes, but in general treat them independently of each other. The specificity of a Marxist theory of banking is in the analytical advantage it could give in tackling this issue. As it will be shown in chapter 5, what according to the bills view broadly defined

appears to be a small change in banking corresponding to a mode of performing banking business in some cases entails substantial changes in the way banks relate to the circuit of capital and hence in the relationship between bank profit and flows of value. Loan sales and securitization would be one example. Thus, the two transformations should not be treated in isolation.

The really interesting question thus is to show that some of the current changes in banking stressed in the literature do not represent as great a transformation as it is often portrayed. Nevertheless, in some cases, this small transformation involves a profound change in the way banking relates to the circuit, hence, in the relationship between bank profit and flows of value – a considerable change that nevertheless remains unnoticed in the literature. Put differently, the transformation of ways of doing the banking business is important, but for reasons other than the ones assumed by the mainstream theorists.

4.4. Marxist Theory of Banking in the Light of Other Theories of Banking

Marx's theory of banking thus understood can serve as a broad theoretical foundation for studying banking and its transformation. It can nest relevant insights of other theories, yet overcome their limitations. Consider how the suggested approach relates to the other existing theories discussed in the previous chapters.

A Marxist theory of banking grounded in the bills view places banks on a foundation other than information asymmetry and market imperfections. This is not to deny that information between borrowers and lenders is asymmetrical – an obvious point, which does not get one far – but rather to acknowledge that even if information were symmetrical there would still have existed a need for banking. In this sense the bills view

has the advantages of the portfolio theories over the information-theoretic approach discussed in chapter 2.

In spite of this similarity, the fundamental differences between a Marxist theory and the portfolio theory should not be overlooked. Contrary to the former grounded in the bills view, the portfolio theory can be viewed as a peculiar mixture of the bills view and the goldsmiths view. In its emphasis on the intermediation function of banking, the portfolio theory resembles the goldsmiths view. At the same time, the theory places a strong emphasis on banks assets and liabilities as promises to pay with different characteristics. Therefore, banks are treated as institutions actively transforming these characteristics. In this sense, the portfolio theories retain an element of the bills view.

This explains the differences between a Marxist theory of banking and the portfolio theories. First, the latter focus on a particular aspect of the bills view – the one when it is applied to the intermediation function of banking. In that sense, the portfolio theories can be regarded as bills view tailored for understanding the realities of banking as it existed in the middle of the 20th century. By the same token, the recent theory of shadow banking as banking (Pozsar et al. 2010) can be viewed as a reinstatement of the bills view to capture the realities of banking today, as portfolio theories captured its intermediation dimension. Contrary to that, a Marxist version of the bills view embraces all forms of exchange of promises to pay, from discounting through intermediation to the modern forms of shadow banking and asset management. Second, a Marxist theory and the portfolio theory differ in their explanation of reasons why the need for this asset transformation arises. For the latter, the reasons can be traced back to differences in subjective preferences of lenders and borrowers. For the former, it is driven by the

objective differences between requirements of value production and the logic of the value leakages from the circuits of capital and revenue.

Another recent theoretical conceptualization of banking was proposed by Gorton (2009), for whom banking is about creating informationally-insensitive debt. This idea fits in a Marxist theory of banking, nevertheless not as an explanation of the *raison d'être* of banks, but rather as an attribute of functioning of banking. Lapavitsas (2003, p. 79-81, 2007, p. 426-427) argues that given that banks do not engage in production and do not create value, establishing trust in their operations is qualitatively different from that for industrial enterprises. Contrary to the latter, credibility of banks' promises to pay cannot be directly assessed based on the economic factors related to value creation or non-economic determinants facilitating their access to liquidity. Hence, there emerges a need for, first, diversification of bank assets to lower the risk of default and, second, an access to reserves. As a result, bank liabilities do obtain a certain degree of independency of individual assets they hold. But there is no insensitivity with respect to information about the bank itself, as the times of bank runs prove. Therefore, informational insensitivity of bank promises to pay objectively arises from the limits in assessing banks' prospects to repay and is based on the trust created and generalized by banks. Nevertheless, there is only partial independency, as bank promises to pay might be insensitive to information regarding individual banks debtors, but not to information about the bank itself.

The relationship between a Marxist theory of banking and theories centered on risk taking and management (Allen, Santomero 1997, 2001) is also apparent. Banking business obviously involves intermediating and reallocating risk for others. But the fact of existence of risk does not provide a basis for arguing that dealing in risk constitutes

raison d'être of banks. Incidentally, this immediate connection between dealing in assets and risk is acknowledged by Allen and Santomero. For them, "by dealing in financial assets, intermediaries are by definition in the financial risk business. By virtue of the fact that they originate, trade, or service financial assets, intermediaries are managing and trading risk" (Allen, Santomero 1997, p. 1478). In this sense Greenspan (1994, p. 3) is right relating all banking activities to some form of dealing with risk. It is inevitable that to the extent that banks deal in assets, they handle risk, as well. But it makes risk handling a secondary, derivative characteristic. Nevertheless, Allen and Santomero mistakenly interpret the ongoing transformation of banking as a change in the fundamental functions of banking, instead of banks finding new ways of making profit based on the same function of exchanging promises to pay.

This brings us to an essential feature of the bills view in general and its Marxist formulation, in particular. Banks are specialists dealing in money, or trading in money (1977 [1894], p. 317, Marx 1991 [1861-63], p. 45-47). Marx stresses that banking functions evolve out of and acquire their content from the functions of money (Marx 1991 [1861-63], p. 42-47). Banks handle risk, because it is an attribute of promises to pay. They also collect information, because it is a necessary counterpart of their business. But they do not emerge as dealers in either risk or information.

A characteristic feature of banks within the capitalist mode of production is their becoming the repositories of temporarily idle funds scattered across the society channeling them into value creation. As a result, money acquires a specifically capitalistic character of loanable capital, which is a two-stage process. First, the function of money as a hoard acquires its capitalist specificity when the circuit of capital comes to

systematically release temporarily idle funds, i.e. when a specifically capitalist form of hoards emerges. Second, these hoards do not remain scattered across the society, but become centralised within the banking system, which gives this money a further determination of loanable money capital. Thus, in a developed capitalist economy banks deal in a specifically capitalist form of money further modified by the functioning of the banking system itself. They deal in loanable money capital.

Incidentally, the emphasis on banks as fundamentally dealers in money is what unites a Marxist theory of banking with the heterodox takes on banking as developed by the circuitists and post-Keynesians. In addition to this common ground, a Marxist theory of banking represents a broader theoretical framework for understanding banking. It incorporates the major insights of the other heterodox traditions, but also overcomes their limitations.

First, as was stressed in the previous chapter, Keynes' major contribution to the theories of banking is his emphasis on liquidity provision. Nevertheless, it is usually neglected that exchange of promises to pay and liquidity provision represent two moments of the same transaction. When a bill is discounted, its seller obtains liquidity and its buyer obtains a promise to pay. When a loan is advanced, the borrower gets an access to liquidity in exchange for a promise to repay with an increment. In the case of banks' deposit taking, bank promise to pay takes the form of guaranteeing liquidity provision on demand or after a certain time period, depending on the form of the deposit. In spite of the two representing the opposite sides of the same coin, it is crucial to stress not only liquidity, as Keynes does, but also the exchange of promises to pay, because the former is a one moment event and does not imply more than just a well known fact of

money being the general form of wealth, or universal equivalent (Marx 2003 [1867], p. 136). The concept of liquidity does not allow one to discuss other social relations involved, in particular, social relations specific to credit. The concept of liquidity and promises to pay operate at different levels of the functions of money. The former sheds light on the social relations at the level of money as coin, i.e. universal equivalent. The latter requires more determinations, as it operates at the level of money as money, namely, money as means of payment and money as hoard. By requiring more determinations, it also sheds light onto a broader set of social relations, including the ones specifically related to credit and not just money. For this reason, for the purposes of analysis of banking it is essential to shift the emphasis away from liquidity to promises to pay, as it would give more mileage for uncovering social relations involved.

Second, as was argued in chapter 3, the circuitist emphasis on the difference between the economic sectors and the need to posit banks in their relationship to these sectors is paramount. Nevertheless, as was also demonstrated before, the way the circuitists locate banks is inherently self-defeating and does not permit to capture the transformation of banking without sacrificing the core of the theory. A Marxist theory of banking allows one to retain the crucial distinction between the sectors and locate banks among them, yet without necessarily reducing it to a single link. Put differently, the theory allows for a multiplicity of links to the circuit. As a result, the transformation of banking can be theorized not as an increase in roundaboutness of productive lending, as in Parguez (2004, p. 46) and Graziani (2003, p. 21), but rather as a change in the prevailing link. This in turn allows one to pose a question of the conditions under which a

prevalent link to the circuit shifts, the social relations underlying this change, and implications thereof.

To recap, a Marxist theory of banking grounded in the bills view yet altered by Marx to account for the specificity of capitalist banking can incorporate the relevant insights of other theoretical approaches to banking by nesting them as particular cases emphasising particular dimensions of banking. Furthermore, it also allows one to overcome the limitation of those theories that have become obvious in the course of the ongoing transformation of banking, hence, a Marxist theory can offer a broader theoretical framework for an analysis of banking and its transformation.

4.5. Bank Profit

As was argued above, the bills view with its emphasis on inherent changeability of banking implies a multiplicity of forms the banking business can take. It would correspond to a multiplicity of the underlying social relations, concrete mechanisms of profit making, and forms this profit would take. Thus, bank profit does not necessarily have to be interest spread, as it is suggested by theories looking at banks as financial intermediaries. Nor does it have to be homogeneous quasi-monopoly rent as asserted by theories emphasising money creation ex nihilo (Bossone 2001a, 2001b, Graziani 2003, Parguez 2004, 2003) and neglecting existence of costs of funds, costly reserves, limits of money creation, and differences in the macroeconomic sources of profit. In this sense the origin of a Marxist theory of banking in the bills view allows for the multiplicity of forms of profit, underlying social relations, and mechanisms of profit extraction, whereas Marx's innovative emphasis on the relationship between banking and the circuits of

capital and revenue gives insights in the macroeconomic sources of bank profit, stressed by a few studies of financialization, but not addressed by theories of banking per se.

4.5.1. Bank Profit in Marx and Hilferding

Given that banks are capitalist firms, one might expect that the general formula of capital would apply to the circuit of banking capital as a particular form of capital. That is, capital invested would abide to $M-C-M'$, with bank profit being simply surplus value.⁶⁷ This impression is reinforced by Marx's discussion of the general formula, when he stresses that it applies to all types of capital, namely, industrial, merchants, and money-dealing (Marx 2003 [1867], p. 153). This take on bank profit might be strengthened further by Sekine who suggests that as long as banks take deposits and make loans, they function similar to merchants who buy cheap and sell dear, with the only difference being that they handle a specific commodity – idle funds (Sekine 1983, p. 422). In that case bank capital would abide to the circuit of merchant's capital ($M-C-M'$), with profit being a form of surplus value.

A direct application of the general formula of capital to the circuit of banking capital would, nevertheless, be misleading. Banks operate in the sphere of circulation, hence, they do not create value and surplus value.⁶⁸ Therefore, the primary concern with respect to bank profit becomes twofold, namely, first, the mechanisms of profit extraction and the grounds of banks partaking in value created in the sphere of production, and

⁶⁷ Probably precisely this desire to fit banking capital into the general formula of capital drives arguments that financial profit comes from surplus value (Fine 2010).

⁶⁸ In this sense bank capital is similar to merchant's capital that also does not create value and surplus value.

second, the macroeconomic sources of this profit. The formal shape of the circuit of capital becomes of secondary importance. This is drastically different from the circuit of industrial capital, because in that case the question of the form of the circuit of capital, mechanism of profit extraction, and macroeconomic sources of profit coincide. Put differently, by arguing that industrial capital abides to $M-C-M'$, one simultaneously addresses the question of the mechanism of profit extraction (buying commodity-labour power) and its source (in the value created by labor power being above its own value). Marx was very clear about it arguing that “propositions about profit, etc., derived directly from the examination of productive capital, cannot be applied directly to mercantile capital” (Marx 1991 [1861-63], p. 64), which for him includes commercial and banking capital.

Given that it is not particularly helpful to try to fit banking capital into the general formula of capital in order to claim that bank profit comes from surplus value, a different approach to bank profit is required. Marx’s discussion of profit in general and banking profit in particular can offer some insights helpful for understanding bank profit. Consider both aspects in greater detail.

The first set of considerations necessary for understanding bank profit originates in the classical political economy discussion of profit in general. In classical political economy, a theory of profit took its clearest form in Ricardo (1821, p. 48-51, 107-130), for whom profit is a part of the fresh flow of value accruing to capitalists as a residual, after workers receive their wages and landlords – rent. Marx adopted this general theoretical framework by treating profit as a part of the fresh flow of value, yet he also fundamentally altered Ricardo’s approach. For Marx, profit is not a residual, but rather an

unpaid part of the newly created value, made possible through exploitation of workers (Marx 2003 [1867], chap. 6-7). The systematic difference between the value of labour power and the value created by workers accounts for existence of surplus value. This take on profit has become the cornerstone of Marx's theory.

Nevertheless, a different approach to profit prevailed before Physiocrats. Profit was argued to arise in the sphere of exchange – a view that according to Marx (1969 [1905-10], p. 41) took its “scientific” expression in Sir James Steuart. For him, profit represents “profit upon alienation” (or “relative profit”) that emerges in the sphere of circulation through redistribution of the already created value – what he calls “a vibration of the balance of wealth between parties” (Steuart 1770a, p. 206). This profit from trade effectively arises through a zero-sum game, when the gain of one of the parties represents a loss for another. Marx argued that it was wrong to identify the capitalist profit with profit arising in the sphere of circulation (Marx 2003 [1867], chap. 5, Marx 1969 [1905-10], chap. 1). Nevertheless, he thought that the concept of “profit upon alienation or expropriation” can offer relevant insights, because “however little it touches the nature and origin of surplus-value itself, [it] remains important in considering the distribution of surplus-value among different classes and among different categories such as profit, interest and rent” (Marx 1969 [1905-10], p. 42). He also used this category in relationship with profit arising in the sphere of financial transactions, especially those involving workers:

[i]nterest may be a mere transfer and need not represent real surplus-value, as, for example, when money is lent to a “spendthrift”, i.e., for consumption. The position may be similar when money is borrowed in order to make payments... In this case interest, like profit upon expropriation, is a fact independent of capitalist production – the production of surplus value. It is in these two forms of money – money as means of purchase of commodities intended for consumption and as

means of payment of debts – that interest, like profit upon expropriation, constitutes a form which, although it is reproduced in capitalist production, is nevertheless independent of it and [represents] a form of interest which belongs to earlier modes of production (Marx 1972 [1905-10], p. 487).

To recap, in the history of economic thought there have existed two approaches to profit – as a fresh flow of value (what Steuart called “positive profit”) and as a redistribution of the already existing value (Steuart’s “relative profit”). Marx identified the former as the major form of profit under the capitalist mode of production, although some elements of the latter can be found in capitalism. This is the first pillar that will be deployed in the present study of bank profits.

The second set of analytical tools comes from Marx’s discussion of the components of banking activity (Marx 1977 [1894], p. 402) that was further developed by Hilferding (1910, chap. 10). Banks operate in the sphere of circulation and in this sense are similar to merchants who, according to Marx, do not create value and surplus value (Marx 1977 [1894], p. 280). Nevertheless, they perform functions indispensable for functioning of the capitalist economies, and on these grounds they partake in value created in the sphere of production. They do it in several ways, hence, receive several forms of profit.

First, they perform the function of money-dealing – monetary services not related to lending. Among them are safekeeping the hoard, settling balances, foreign exchange operations, and book-keeping (Marx 1977 [1894], chap. 19). These are the functions of the total social capital in the sphere of circulation⁶⁹, and banks become specialists in

⁶⁹ In Volume 2 of *Capital* Marx (1967 [1885], p. 48) discusses functional forms taken on by capital-value at the various stages of its circulation. He distinguishes between money-capital, commodity-capital, and productive capital. Thus, money-dealing capital emerges when the functions of money-capital get to be performed by a group of capitalists, specializing in these operations (Marx 1977 [1894], p. 315).

performing them on behalf of the total social capital, hence, do it more efficiently and minimize the pure costs of circulation. On these grounds banks partake in the value created in the sphere of production and earn an average rate of profit. This profit is a deduction from the surplus value (Marx 1977 [1894], p. 322).

Second, banks engage in advancing loanable capital, and earn interest for parting with liquidity for a specific period time. The interest earned falls into three parts, namely, interest paid on deposits, costs of doing banking business (both constant and variable capital), and bank profit. Focusing mainly on bank lending to industrial and commercial capitalists, Marx considers this interest as a deduction from surplus value created in the sphere of production (Marx 1977 [1894], p. 378-381). Nevertheless, lending operations of banks, like any lending activity, does not earn average rate of profit (Marx 1977 [1894], p. 362-364). The interest rate is determined merely by the supply of and the demand for loanable capital and has no immediate relationship to the rate of profit.

An important aspect of Marx's treatment of banking is that both components of its profit come from surplus value created in the sphere of production. Thus, for Marx, bank profit is a part of the fresh flow of value. This approach is further developed by Hilferding who analyzes two more forms of bank profit.

First, for Hilferding, banks come to perform the underwriting function by floating shares. On these grounds they earn a part of the founder's profit (Gründergewinn⁷⁰)

⁷⁰ The German category Gründergewinn was translated by Watnick and Gordon in the Bottomore edition of *Finance Capital* as "promoter's profit". Itoh (1988, p. 404-405) suggested a more accurate translation might be "founder's gain". It is a more literal translation of the original German term. Moreover, this type of gain can accrue to founders of an enterprise when it is converted into a joint-stock company. Founder's profit therefore exists regardless of incorporation being facilitated by a promoter (an individual or a bank). Given that the economic content of this form of gain stems from the process of incorporation per se, regardless of who facilitates underwriting and distribution of the shares, in the present study Hilferding's concept Gründergewinn will be referred to as "founder's profit".

which for Hilferding is “an economic category *sui generis*” (Hilferding 1910, p. 112, italics in the original). It is a special kind of profit because it is an economic form taken by a new type of social relations – separation of management and ownership and a related genesis of the joint-stock capital. Following Marx, Hilferding assumes the interest rate to be below the rate of profit in the sphere of production (Marx 1977 [1894], p. 358, Hilferding 1910, p. 109-110). Given that buying shares is formally akin to investing loanable capital, he further assumes the interest rate to be equal to the dividend yield. It is worth noting that Hilferding is not entirely consistent on this, and at times he mentions the risk premium which would bring the dividend yield above the interest rate (Hilferding 1910, p. 108). Nevertheless, in the final analysis, trying to map Marx’s distinction between interest and profit of enterprise onto the difference between dividends and founder’s profit, respectively, Hilferding assumes the risk premium away. As a result, he argues that the founder discounts the future surplus value (S) with the rate of profit (r), whereas shareholders – with the interest rate (i). Given that the interest rate lies below the profit rate ($i < r$), the flow of the surplus value discounted by the former exceeds the surplus value discounted by the latter ($\frac{S}{r} < \frac{S}{i}$). Thus, the differences in the rates of return bring the price of the shares over and above the amount of capital invested. The difference between the two constitutes founder’s profit which for Hilferding is profit of enterprise accruing in a lump sum (Hilferding 1910, p. 129). Given that banks facilitate floating the shares, they receive a part of this profit.

Second, banks engage in trading securities – an activity identified by Hilferding as speculation. As a result, they earn what he calls speculative profit – the other, fourth, type of profit. For Hilferding (1910, p. 135), “speculators gain only from each other.

One's loss is the other's gain”, so speculative profit is a zero-sum game among the financial institutions. In other words, it is what Steuart called “profit upon alienation”.

Hilferding obviously advanced Marx’s theory of banking and bank profit. Nevertheless, his treatment of bank profit is problematic at times. First, as noticed by Itoh (1988, p. 287), Hilferding’s claim that founder’s profit is a profit of enterprise accruing in a lump sum is not convincing, for a number of reasons. The most important one is that if founder’s profit is a capitalized future, i.e. not yet created, surplus value, its macroeconomic source remains unclear. How can one partake in value that does not exist yet? Furthermore, the exact mechanism of how this not yet existing surplus value can be appropriated is also not explained by Hilferding. By the same token, he does not pose the question of the macroeconomic consequences of the existence of this type of profit for distribution of national income, price dynamics, and so on. Second, his treatment of the so-called speculative profit is even less successful. By labeling it a zero-sum game, Hilferding manages to stay away from the question of the relationship between trading gains and flows of value. Nevertheless, as it will be shown below, a relationship between the two does exist, although it is not apparent and mediated. Certain transactions associated with capital markets, including trade in assets, can be positive-sum games involving wealth transfers due to differences in the rates of return. These gains can be further magnified by fraud and speculative bubbles.

To recap, the emphasis on the distinct forms of bank profit, with each of them associated with specific social relations, mechanisms of profit extraction, and different connection to the flows of value, is a hallmark of Marx-Hilferding theory of bank profit. The present analysis of bank profit and its transformation will be based on both

cornerstones of this theory. Specifically, we will deploy Marx's distinction between profit as a part of the fresh flows of value, and profit as a redistribution of the already existing flows of value. We will also start by acknowledging that bank profit should be viewed as an envelope category comprising a number of forms of profit, each having distinct underlying social relations, and mechanisms of profit extraction. Bank profit should not be treated as interest spread, as it would follow from theories based on banks as intermediaries, nor can it be treated as a homogeneous quasi-rent as asserted by theories emphasizing money creation *ex nihilo* (Bossone 2001a, 2001b, Graziani 2003, Parguez 2004, 2003). Furthermore, the relationship between the two moments of Marx's theory of profit, namely, its relationship to the flows of value and the multiplicity of forms, are not always apparent in his writings. It is especially clear in the case of bank lending to workers, when profit is interest spread, yet is not a part of surplus value. Therefore, we will bring them together by stressing how different forms of bank profit are connected to flows of value. This will be an organizing principle of what follows. It is worth noting that this approach would allow us to incorporate the observation by DeYoung and Rice (2004a) that the composition of bank revenues is important and, even more, this composition is the most appropriate lens for studying the transformation of banking. At the same time, the general analytical framework outlined above would allow us to go beyond DeYoung and Rice's empirically-driven analysis, and uncover both the social relations corresponding to these types of profit and their macroeconomic sources. The emphasis on the macroeconomic sources from a political economy perspective would be in line with Pollin's (1996) insightful argument discussed in chapter 3.

4.5.2. Bank Profit and Loanable Capital

To develop an analysis of bank profit further, we need another analytical tool that is grounded in Marx's and Hilferding's discussion of finance, yet goes beyond them by incorporating insights of Marxists who worked on these questions later. The concept of loanable capital is decisive for these purposes.

Marx defines loanable capital as capital available for lending, i.e. "intended for bearing interest" (Marx 1977 [1894], p. 478)⁷¹. It is true though that he does not start his analysis of finance with this concept. His starting point is the category of interest-bearing capital, which, as was briefly mentioned above, has two specifically capitalist dimensions. Consider it in a greater detail.

First, and widely accepted, interest-bearing capital is a credit relationship corresponding to borrowing for the purpose of embarking on a circuit of capital (Marx 1977 [1894], chap. 21, 23). The second, and less appreciated specifically capitalist dimension of interest-bearing capital, is its origin in temporarily idle hoards generated by the circuit of capital itself. As shown by Marx (Marx 1977 [1894], p. 403, Marx 1967 [1885], p. 78, 83-85, 321-322, 451, Marx 1991 [1861-63], p. 169), further developed by Itoh and Lapavitsas (1999, p. 65-69), and more formally by Lapavitsas (2000), and also emphasized by Campbell (Campbell 2002, p. 215), the circuit of capital systematically releases temporarily idle hoards, "leakages of value" that are collected by the financial sector and converted into interest-bearing capital. As a result, the circuit of capital itself

⁷¹ Loanable capital is money capital concentrated within the financial system, thus, converted from passive (potential) to active capital. "One can understand the pleasure experienced when all these potential capitals within the credit system, by their concentration in the hands of banks, etc., become disposable, "loanable capital," money-capital, which indeed is no longer passive and music of the future, but active capital growing rank" (Marx 1967 [1885], p. 493).

forms a foundation for a systematic hoard formation, hence, for functioning of the credit system. This is further reinforced by the financial system coming to accumulate temporarily idle hoards scattered across the society as a whole, thus, the circuit of revenue becomes the other foundation for a systematic hoard formation. As a result, the financial system comes to be the repository of the temporarily idle funds that are placed for the needs of capital accumulation. As Marx put it, “borrowing and lending money becomes their [money-dealers’] particular business... this aspect of the banking business consists of concentrating large amounts of the loanable money-capital in the bankers’ hands... They become the general managers of money-capital” (Marx 1977 [1894], p. 402, also p. 368). The fact that the financial system assembles funds across the society is also the basis for Harvey’s conclusion that in the financial system the circuits of capital and revenue are mixed together (Harvey 2006, p. 273-276), so that Marx’s careful distinction between the money-form of revenue and the money-form of capital (Marx 1977 [1894], p. 443) is abandoned in practice by the very functioning of the financial system. It essentially means that loanable capital is homogenized by the financial system.

Loanable capital is a more developed form of the interest-bearing capital, both historically and logically. Therefore it is not surprising that Marx extensively uses the former category in his analysis of banking at a lower level of abstraction (Marx 1977 [1894], chap. 30-32).

First, the very functioning of the financial system mitigates the importance of the purpose of borrowing. From a bank’s perspective, it does not matter whether a loan is extended for a productive purpose, or for a purpose of household consumption – an observation which, as was shown in chapter 2, becomes a foundation of all the

mainstream theories of banking, for which the difference across economic sectors does not exist. Incidentally, this banks' indifference to the purpose of borrowing is entirely consistent with Marx's methodology. He argues that the quantitative division of profit into interest and profit of enterprise leads to a qualitative division of capital into capital-as-property and capital-as-function (Marx 1977 [1894], p. 372-379). Thus, an ability of bearing interest becomes viewed in the society as a capacity of any sum of value. A logical conclusion is that the origin of interest in surplus value becomes unrecognizable, and the link between lending and the necessity for surplus value production as its material support – if this lending were to have a systematic self-sustained foundation – becomes obscured. Therefore, it is not surprising that the purpose of borrowing ceases to be a decisive feature of the credit relationship.

This is reinforced by the fact that credit relations pre-date capitalism: there is nothing unprecedented in borrowing for purposes other than starting a circuit of capital. Nevertheless, there is also a fundamental difference between this indifference to the purpose of borrowing on a capitalist basis and within the previous modes of production. If usury was condemned in the feudal society as parasitic and predatory, the qualitative division of capital into capital-as-property and capital-as-function rationalizes what used to be condemned, and gives essentially usurious practices social justification which it could never have before. Thus, the development of capitalism rationalises, justifies, and subsumes certain pre-capitalist relations. While usurious practices used to be condemned, they are nowadays viewed as an entirely justified profit-making activity of modern banks. And this happens in spite of this activity having taken a grander scale, with a larger share

of population absorbed into it, and its being based on deploying idle funds of the society as a whole, and not just own capital of an individual usurer.

Second, as the banking system develops, banks come to create money, so that their reliance on the results of previous accumulation and idle hoards scattered across the society is also mitigated. Thus, loanable capital originates in temporarily idle hoards, but also in a pure money form created by the banking system itself. An important aspect here is that in both cases loanable capital has a purely social nature. In the former case, it is amassed idle funds of the society as a whole (Marx 1977 [1894], p. 366), in the latter case, money creation by banks is based on acceptability of demand deposits as means of circulation, i.e. on deposits acquiring money-like features. Furthermore, more importantly, in both cases there is no immediate relationship between loanable capital and flows of value.

When loanable capital originates in the idle hoards, these hoards either stem from fresh flows of value or represent hoarded leakages of value coming from the previous cycles of accumulation. In spite of the immediate origin in the flows of value, once these hoards are amassed by the credit system, they leave the circuit of capital and revenue, and enter a separate circuit – that of loanable capital. The lack of connection between loanable capital and flows of value is even more apparent when banks create their own money. In this case loanable capital originates within the financial system itself. An important point here is that loanable capital has its own circuit as distinct from the circuits of capital and revenue, and no matter what is the origin of its elements – hoards scattered across the society or bank money – once they become a part of the circuit of loanable capital, they cannot be a part of the circuit of revenue or capital at the same

time. To be a part of one of these two circuits, money would need to exit the circuit of loanable capital. This argument is somewhat related to Keynes' distinction between industrial and financial circulation (Keynes 1971 [1930], chap. 15).

An important characteristic of loanable capital is its inherent tradability. Discounting bills of exchange is the simplest form of a trade in securities. This inherent tradability is accentuated by the very functioning of the financial system. On the one hand, it reinforces the possibility of trade in loanable capital by extending the range of counterparties engaged in this trade, expanding variety of financial instruments, and homogenizing the methods of trade. On the other hand, the financial system also strengthens the need for this trade. Issuing short-term liabilities characteristic to banking and specializing in lending on demand, as in the case of some types of banks (Smith-like banks, in line with the Anglo-Saxon "sound banking" tradition), makes banks the main source of demand for liquid funds. This demand for liquidity by banks themselves is further reinforced by the need to settle inter-bank obligations, meet reserve requirements, and by banks increasingly issuing contingent liabilities. As a result, tradability of loanable capital increases. The inherent tradability of loanable capital is important because it shows that it is misleading to treat the trade in financial assets as an arbitrary activity of the financial sector, even less as pure gambling, or speculation.

The discussion of loanable capital is of paramount importance. It will be shown in the following section that some characteristics of this form of capital are essential for an analysis of the bank profit. Among those are inherent tradability and existence of a separate circuit of loanable capital that does not have an immediate connection to the circuits of flows of value, be it the circuit of capital or the circuit of revenue.

4.5.3. Bank Profit: Beyond Marx and Hilferding

It was argued above that the content of all financial transactions can be understood as an exchange of loanable money capital for a promise to pay. Banks are specialists dealing in loanable capital. Therefore, in the most general terms, bank profit arises from this exchange of loanable capital for a promise to pay.⁷²

There are two broad types of profit associated with promises to pay. This major distinction will lie at heart of the subsequent analysis. First, there are profits arising from *holding* promises to pay issued by others. These profits are justified by parting with money for a specific time period, and normally take the form of interest payments, coupons or dividends. Those come directly from flows of value – surplus value (profit) in the case of the promises being issued by capitalists or value of labor power (wages) in the case of promises issued by workers. These profits represent a direct deduction from these flows, because, in the former case, the money form of capital is essential for starting a circuit, therefore, the credit relationship helps generate the source of repayment of interest. Interest from lending to workers comes directly from their wages, as it enters the value of the labour power together with the value of the commodities that involve borrowing and paying interest. In this sense, although Harris (1976, p. 160-161) mistakenly treated interest paid by workers as profit on merchant's capital, thus conflating interest with merchant's profit, he was right in his emphasis that the value of

⁷² The only exception is profit from money-dealing discussed above.

labor power comprises both interest on borrowing in order to purchase specific commodities and the value of commodities themselves (cash price).⁷³

Nevertheless, profit associated with holding promises to pay is not the only form of profit related to the promises. The second form arises from various ways of *dealing in, or handling* these promises. Its most obvious example is trading revenue, both on own account (proprietary trading) and fees from trading for others. But, as it will be shown in the next chapter, more complex ways of dealing in promises to pay have similar characteristics. These other ways of handling promises result in profit from underwriting (and securitization income and mergers and acquisitions fees as its particular form), and to some extent even asset management fees.

There are essential differences between these two types of profit. First, profit from holding promises accrues to the one who parts with liquidity and obtains a promise. By contrast, profit from handling promises accrues to the one parting with a promise and restoring the money form of capital advanced. Therefore, liquidity provision, which is a social justification of profit in the former case, cannot justify profit from dealing in promises. Second, profit from holding promises presupposes that money leaves the circuit of loanable capital and enters the circuit of capital or revenue. On the contrary, profit from dealing in promises does not have this prerequisite, and it accrues to the dealers in loanable capital as money travels within the circuit of loanable capital itself.

Third, as it will be illustrated below, these two differences result in different

⁷³ If one adopts a definition of value of labor power from volume I of *Capital*, interest payments by workers would imply that wage is greater than the value of labor power. It leads some to conclude that interest payments by workers come from surplus value (see, e.g., Fine 2010). The present study assumes that profit in the sphere of circulation is better theorized as a combination profit from redivision of surplus value (as in the case of merchant's profit, profit from money dealing, and interest from lending to capitalists) and "profit upon alienation" (unrelated to redivision of surplus value). This argument is elaborated in Lapavitsas and Levina (2011).

macroeconomic sources of profit. Profit from holding promises involves loanable capital leaving its circuit, and entering a circuit of capital or revenue, therefore, associated profit can be drawn from one of these circuits. By contrast, profit from dealing in promises can accrue without money leaving the circuit of loanable capital, therefore, this profit does not have an immediate relationship with flows of value. Rather, it can come from loanable capital itself.⁷⁴

If the sources and mechanisms of profit extraction from holding promises to pay are straightforward and represent re-division of the flows of value, profit from dealing in promises deserves a more careful analysis. It was argued in section 4.2 that the simplest form of credit relations can be identified with commercial (trade) credit. By the same token, the embryo form of dealing in promises to pay can be traced back to discounting the instruments of trade credit. In spite of the apparent simplicity of bills of exchange and related transactions, some relevant insights on the nature of profit from dealing in loanable capital can be gained from an analysis of these simple instruments.

Assume that a bill of exchange is generated through the sale of commodity output, and is then discounted by other capitalists. The property that matters for our purposes is that the ultimate holder of the security is not the sole receiver of interest payments. Rather, the payments made by the issuer of the security are distributed among all participating capitalists, depending on the period of holding the bill as well as the rate of

⁷⁴ Under certain circumstances profit from holding promises to pay can also come from loanable capital. This would happen when interest or dividends are paid by means of borrowing. Following Minsky (2008 [1986], p. 231), this case is normally referred to as Ponzi finance. Nevertheless, this is a characteristic of pre-crisis periods and cannot be viewed as a general case. Therefore, profit from holding promises to pay can be argued to come from loanable capital only in exceptional cases, whereas it is the general rule in the case of profit from dealing in promises to pay.

interest (Itoh, Lapavitsas 1999, p. 94, Itoh 1988, p. 266).⁷⁵ This insight can be elaborated through a simple numerical example.

Assume capitalist A buys commodities from B, with cash price of \$100, issuing a bill of exchange amounting to \$110 due in 2 months. Suppose now that B sells the bill to capitalist C for \$105. When the bill matures, there can be two broad analytical outcomes.

The first is that A generates the expected flow of surplus value, which is then distributed among the bill holders, mediated by a transfer of loanable capital.⁷⁶ Thus, when the bill falls due, C receives \$110 from A, and in effect the total interest of \$10 is equally distributed between B and C. In this case, B's gain has a dual nature. Ultimately, it is part of the interest paid by A, originating in the surplus value created by A. But in immediate terms, it comes directly from the loanable capital of C advanced in exchange for the bill. Note that it is the tradability of the debt instrument that opens up the possibility of inter-temporal distribution of flows of value.

The other possible outcome is a pure redistribution of the loanable capital of the ultimate holder. This would occur if A failed to generate the expected surplus value and paid only p_a when the bill fell due. Three analytically important cases are evident (Figure 1).

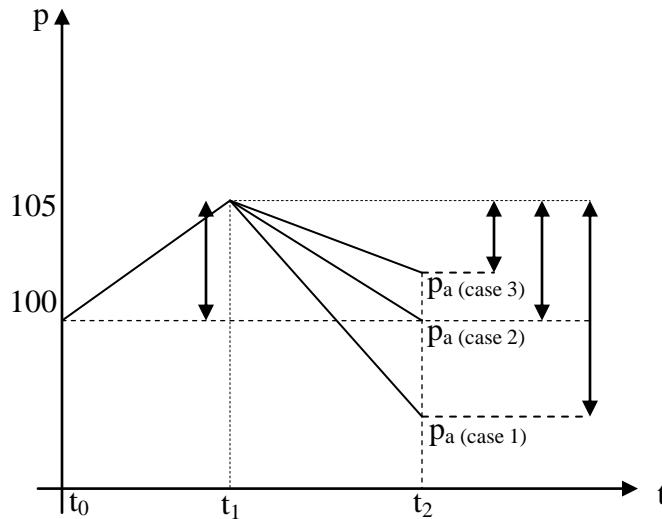
⁷⁵ For commercial credit, interest is the difference between the credit price and the cash price. This severely constrains analysis of the impact of interest rates on financial profits, even though some conclusions could still be drawn. For this reason, it is best to examine the role of interest rates in the context of bonds, which is also the natural terrain for analyzing the relationship between interest rates and capital gains. Henceforth, for trade credit instruments, discounting at a price higher than the cash price will be interpreted as a result of approaching maturity. Nevertheless, it is also conceivable that a promissory note could be immediately resold at a price higher than it was bought. This would be analogous to a change in the interest rate, thus providing an explicit link between, on the one hand, simple trade credit and, on the other, bonds and equities.

⁷⁶ Formally speaking it should be 'idle funds' and not 'loanable capital' of C since the category of loanable capital simply does not exist at the level of abstraction of commercial credit among capitalists. But this would merely complicate expressions without offering any analytical benefits.

Case 1: $p_a < 100$. The total loss of B and C would be equal to $100 - p_a$. This would be less than the loss of the ultimate holder, C, which would amount to $105 - p_a$, arising from A's underpayment but also from reallocating C's loanable capital in favor of B. By the same token, there would be a gain of \$5 for B originating in the appropriation of C's loanable capital.

Case 2: $p_a = 100$. Total losses of B and C would be equal to zero, but B would make a gain of \$5 from the redistribution of C's loanable capital. This is a pure zero-sum game.

Figure 1: Distribution of gains from trade in a bill of exchange



Case 3: $100 < p_a < 105$. Total gains of B and C would be equal to $p_a - 100$, with the ultimate source of these gains being surplus value produced by A. Nevertheless, C would still have losses amounting to $105 - p_a$. This implies that B's gains are the sum of some of C's loanable capital and some of the flow of surplus value. Thus, even when a flow of surplus value is produced and distributed, there exists the possibility that profit would exceed surplus value by appropriating a part of the loanable capital of another capitalist. In immediate terms B's gain arises from C's loanable capital. Once mediated

by the (inadequate) return of loanable capital to C, some of B's gain would arise from the flow of surplus value, but some would still be a part of C's loanable capital.

All three cases represent the redistribution of the ultimate holder's loanable capital in favor of the intermediate capitalist, B, and each represents a different degree of total loss.⁷⁷ The gain from selling the bill comes either completely from C's loanable capital, or from a combination of C's loanable capital and the surplus value created by A.⁷⁸

There are several relevant conclusions from this analysis of a trade in instruments of commercial credit. First, there is no immediate relationship between profit from discounting a bill of exchange and flows of value. This profit accrues apparently independent of and prior to the value generating process. Second, nevertheless, this does not imply that there is no connection between this type of financial profit and flows of value. Profit from discounting a bill has a dual nature. Immediately, it comes directly from loanable capital of the bill buyer.⁷⁹ The subsequent return of loanable capital to the ultimate bill holder mediates the content of profit of the bill seller. Therefore, once mediated, this profit ultimately becomes associated with the future flows of value. In this case the re-division of loanable capital facilitates an inter-temporal redistribution of flows of value. Put simply, the security seller gets profit without corresponding output

⁷⁷ B's gain is a type of appropriation through exchange where the market value of the item exchanged is not known at the moment of transaction. The gain would come from either good luck or fraud on B's part.

⁷⁸ A word of caution is in order. Joint liability characteristic of commercial bills could limit the reallocation of C's loanable capital in favor of B, thus modifying the redistributive results.

⁷⁹ Incidentally, loanable money capital as an immediate source of capital gains was identified by Marx in his analysis of cyclical fluctuations in the capital markets and related fluctuations in interest rates. He argued that, if securities are bought when interest rates are high, hence stock prices low, and sold when stock prices regain their level, "a portion of the money-capital of the public is thus appropriated" (Marx 1977 [1894], p. 502).

produced, and the ultimate holder gets the whole surplus produced, but a part of it is used to compensate for the monetary wealth transferred to the previous holder of the security. If mediation is incomplete, i.e. if returns fail to materialize as expected, profit from discounting represents a pure redistribution of loanable capital. Thus, it is a reinstatement of “profit upon alienation”, but with respect to loanable capital.

The dual nature of profit from handling loanable capital has important implications.

First, this profit obtains its definite content only upon mediation. Depending on whether loanable capital (and its increment) returns to the ultimate holder of the promise to pay, the profit of the intermediate dealers becomes either a part of the flows of value inter-temporally distributed, or profit upon alienation when loanable capital is re-divided. But until loanable capital is recovered by the ultimate holder of the promise, profit from dealing in promises remains a “pure form”. This is what might be called “fictitious profit”.⁸⁰ It is fictitious in the sense that it comes from loanable capital itself, which, as was argued above, has its own circuit. To some extent loanable capital originates in the flows of value, but the rest of it originates through issuing promises to pay that are not related to flows of value. Individual constituents of this mass of loanable capital might or might not have a counterpart in the freshly produced value. Therefore, taken as a whole, loanable capital does not have a counterpart in the newly created value, and profit coming from loanable capital also cannot be said to correspond to the current flows of value.

⁸⁰ After the 2007-2009 crisis, there have emerged (on the left but not only) some ideas of “fictitiousness” or “falseness” of financial profits, but those were just words and, at most, good intuitions, without any content, let alone a theoretical analysis. In the present study this “fictitiousness” of profits coming from re-division of loanable capital does not mean that the profit is somehow not real, it only means there might be no corresponding value produced at the time of accrual of this profit.

Incidentally, the fact that profit from re-division of loanable capital originates outside of the flow of capital and revenues is acknowledged by National Accounts that exclude capital gains from current national output.

Second, in spite of not having a counterpart in the newly created value, profit coming from re-division of loanable capital acquires existence on a par with profit from the fresh flows of value. Put differently, this “fictitious” profit has real effects on the economy. Prior to the time of final repayment, the final holder can borrow against the expected market value of the bill and spend what is borrowed. Furthermore, a realized gain made by the seller of a bill, which comes from the ultimate holder’s loanable capital, is indistinguishable from profit backed by flows of value. This profit can be spent at any point of time, prior to or after the maturity of the bill. This becomes possible due to a peculiar role of money in capitalist economies. Profit coming from loanable capital, like any other profit, has a monetary form. Money is primarily a universal equivalent, i.e. is in a monopolistic position of direct exchangeability with other commodities. It is the absolute form of existence of exchange value, or the universal commodity (Marx 2003 [1867], p. 136). Hence, although some parts of profit from dealing in loanable capital might not have flows of value as their counterpart, it does not prevent this profit from having the same purchasing power as profit backed by flows of value. Therefore, this profit can have redistributive impact on the current output.⁸¹ Furthermore, in addition to the monetary form of this profit, the other factor contributing to this profit having existence and real effects on a par with profit backed by current flows of value is the wide-spread view that whatever earns profit corresponds to useful services. An example

⁸¹ Profit from re-division of loanable capital has two real effects on the economy. It affects, first, the present distribution of output and, second, future monetary wealth of the asset holders.

of this approach can be found in Samolyk (2004, p. 30, 55) for whom profit reflects the amount of financial services provided. She does not question whether some of these profits can exist without associated useful services.

Third, the ability of the financial system to create its promises to pay that are universally accepted as means of circulation places it in a strategically important and privileged position compared to the other sectors. By augmenting loanable capital beyond the idle hoards leaking from the circuits of capital and revenue, the financial system can create loanable capital, hence, the source of its own profit. This loanable capital would need to be validated by the circuit of capital through generating the hoards ex-post, nevertheless, in the short run the financial system can create the source of its own profit.

Therefore, fourth, there is an intrinsic motive for financial institutions to generate expansion of loanable capital, or, what is usually called “overextension in balance sheets” (Borio 2007). It is not surprising that financial institutions feed on bubbles, and periods of euphoria are endogenous to this type of profit.

Finally, given that profit from re-division of loanable capital in immediate terms comes from loanable capital itself, and not from fresh flows of value, at each point in time there is no immediate relationship between the magnitude of this profit and total output produced. Put differently, this profit is not immediately related to and, therefore, not limited by the value of the current output, and even less so by the current surplus value.

4.6. Conclusion

The chapter argues that in the history of economic thought there have existed two broad approaches to banking – the bills view and the goldsmiths view. Marx has adopted

the general theoretical framework of the former by, first, tracing the origin of bank credit back to commercial credit and treating banking as a business of exchanging promises to pay, and second, by emphasizing the active principle behind the functioning of banks that strive to increase their lending capacity. At the same time, Marx has fundamentally altered the ahistorical bills view by explicitly locating banks within the capitalist mode of production.

This dual foundation of Marx's theory of banking gives it an analytical advantage in theorizing banks, in general, and in addressing their ongoing transformation, in particular. Contrary to the theories discussed in the previous chapters, Marx's theory of banking can capture both a change in the mode of performing banking business and a change in the character of the borrower.

In particular, it is argued that contrary to the wide-spread arguments, the ongoing transformation of the ways in which the banking business is performed is not that unprecedented. Deposit taking and issuing bank notes can be viewed as the original forms of financial innovation, and liability management and asset sales are merely new ways of banks raising their lending capacity. At the same time, the other aspect of the transformation of banking – its turn to households – is shown to be a more profound change in a broader historical context. It represents more than just a reinstatement of usury. On the one hand, contrary to usurers, banks use for essentially usurious activities not only their own capital, but the truly social funds scattered across the society and money created by banks themselves on the basis of trust bestowed on them by the society. On the other hand, paradoxically, banking becoming a capitalist enterprise like any other gives these usurious activities a rational justification. As a result, what used to

be condemned in the previous modes of production attains rationalization, in spite of taking even worse forms and a larger scale.

As a result, a Marxist theory of banking is shown to be able to nest relevant insights from other theories, while also overcoming their limitations.

This discussion of a Marxist theory of banking serves as a foundation for understanding bank profit. The present study deploys two relevant analytical considerations coming from the classical political economy, Marx, and Hilferding. First, in general, profit can originate in fresh or in existing flows of value. Second, as stressed by Marx and Hilferding, banking business is constituted by multiple activities each earning a distinct type of profit.

In spite of the study being explicitly rooted in the Marx-Hilferding theory of bank profit, their theory requires a further development. Two analytical considerations become paramount in accomplishing it.

First, the concept of loanable capital is important. On the basis of inherent tradability of loanable capital and its having its own circuit, it is shown to be a separate source of profit, in addition to flows of value.

Second, it is useful to distinguish between two types of profit associated with promises to pay – profit from holding them and profit from dealing in them. The former is straightforward, as this type of profit has had enough attention before and is shown to originate in flows of value, fresh or existing. The latter has not been studied at a theoretical level before and represents a more complex case.

Profit from dealing in promises to pay is shown to have a dual nature. In its immediacy, it comes from re-division of loanable capital. Once mediated by a return of

loanable capital of the ultimate holder of the promises to pay, this profit becomes a part of surplus value. If the mediation is incomplete, this profit remains a pure redistribution of loanable capital. The former represents an inter-temporal redistribution of flows of value mediated by reallocation of loanable capital, the latter is a reinstatement of “profit upon alienation”.

It was argued above that there are many forms of profit from non-lending activities associated with the transformation of banking. Among them are profit from trading financial assets (both on own account and for others), underwriting revenues, fees for mergers and acquisitions, securitization revenue, and asset management fees. Therefore, a question arises how these forms of profit can be explained on the basis of the suggested theoretical framework. This will be an object of study in the next chapter.

CHAPTER 5

POLITICAL ECONOMY OF CAPITAL GAINS: THROUGH THE LENS OF TWO SETS OF PRICES OF CAPITAL GOODS

5.1. Introduction

The goal of this chapter is to argue the rise of the financial markets brings about a particular form of profit making that restores the predatory aspects of finance by raising them on a capitalist foundation. To make this point, the chapter uses the dual price system of capital goods – the market price based on capitalization of their future yields and the replacement costs – to show the gap between the two prices creates a foundation for extraction of a particular type of financial profit that rose to prominence over the past few decades. A starting point of analysis of this form of profit should not be speculation. Instead, the gain is shown to come from redistribution of monetary assets scattered across the society, be these already existing or newly created assets. This type of financial profit can be extracted through a variety of mechanisms, with an initial public offering being the simplest of them. Among the other mechanisms are mergers and acquisitions and securitization. In all these activities liquidity is necessary for this form of profit making, therefore, banks perform a necessary function of liquidity provision which in turn allows them to share in financial profit associated with these transactions. Given that the non-traditional banking activities involve a more profound transformation of bank revenues and profits than of the bank function which remains to be liquidity provision, a fuller understanding of the social significance of the transformation of banking requires using the character of bank revenues as the lens for analysis. This argument is developed as follows.

Section 2 discusses the dual price system of the capital goods as a basis for a theory of investment introduced by Keynes and further developed by Tobin and Minsky. These theories, however, do not explain a number of empirical facts due to conceptual problems with these approaches to investment. Drawing on Mehrling and Hilferding, section 3 argues that the problematic character of these theories stems not from the distinct conceptualization of capital goods as having two sets of prices but rather from an attempt to use it as a general theory of investment. The dual price system is argued to be relevant, as it signifies a transformation of the concept of capital in the 20th century. The social significance of this transformation lies in a rise of a particular form of profit making associated with a rise of the capital markets. Section 4 compares Hilferding's and Minsky's approaches to profit associated with the capital markets and argues that they agree that this form of profit arises due to the gap between the two sets of prices, that a starting point of analysis is not speculation but rather realization of gains associated with value that is not yet created. Hilferding, however, differs from Minsky in identifying the difference in the rates of return as the social foundation of extraction of this form of profit, whereas for Minsky the capital markets remain a black box that merely changes a mode of valuation. It is argued in section 5 that the financial profit associated with the gap between the two sets of prices requires specifically capitalist social relations, yet simultaneously reinstates the pre-capitalist, predatory forms of profit making. Section 6 shows this profit comes in many forms and accrues not only to financial institutions, but also to non-financial institutions and individuals, making the usual criticism of the financial system too narrow. Section 7 focuses on the concrete forms taken by profit due to the gap between the two sets of prices and on this basis reevaluates the question of the

transformation of banking. Initial public offerings, mergers and acquisitions, and securitization are suggested to be analyzed as concrete mechanisms of profit making. These transactions require liquidity making banks well suited to facilitate these transactions, which in turn becomes the basis for banks' sharing in profit associated with the dual price system. On these grounds it is concluded the significance of the transformation of banking cannot be fully recognized on the basis of bank functions, and the focus of analysis should rather be shifted to the character of bank revenues and profits. Section 8 concludes.

5.2. From the Dual Price System to a Theory of Investment: Keynes, Tobin and Minsky

Keynes offered a distinct conceptualization of capital goods that lies at the heart of his theory of investment further developed by Tobin and Minsky. The debate around the theoretical and empirical validity of these theories offers, however, good reasons to believe that Keynes-Tobin-Minsky approach to investment has a limited explanatory power, and the gap between the market value of capital goods and their replacement costs cannot be a foundation of a general theory of investment.

Keynes (1973 [1936], p. 135-137) suggested that capital goods – what he called “capital-assets”, or investment – have two sets of prices, namely, the demand price and the supply price. The demand price of the investment is determined by the prospective yield from an asset and the current rate of interest. The prospective yield is a difference between the selling price of the goods produced with an aid of the capital asset and the running expenses during the life of the asset. The supply price of the capital-asset is its

replacement cost. Keynes used this conceptualization as a starting point of his theory of investment. For him, investment is carried out up to a point when the supply price of investment is equal to its demand price.⁸²

This approach was subsequently formalized by Brainard and Tobin (1968) as a q-theory of investment. As in Keynes, for Brainard and Tobin (1968 p. 103-104), the key determinant of investment is the relationship between the market valuation of equities and the replacement cost of the physical assets they represent. Investment is stimulated when the market valuation exceeds the replacement costs, which is equivalent to the yield on equity being below the return to physical investment.

This conceptualization of physical capital as having two sets of prices has entered into a general equilibrium approach developed by Tobin (1969). The physical capital at reproduction costs (p) generates profits at a rate equal to the technologically determined marginal efficiency of capital (R). Tobin allows the market value of the physical capital to diverge from its reproduction costs due to the required rate of return for investors diverging from the marginal efficiency of capital. If the market value of capital is qp, the rate of return for investors is $r_K = R/q$. Tobin emphasized the divergence of the two sets of prices in the short run (Tobin 1969, p. 29)⁸³. A long run equilibrium condition, however, requires the market value and the reproduction costs of the physical capital to equalize.

⁸² At this point, the current rate of interest is equal to the marginal efficiency of capital (MEC), as Keynes defined the MEC as the discount rate making the present value of the stream of future net revenues from use of the asset equal to its supply price. Marginal efficiency of capital in Keynes is equivalent to what Irving Fisher (1930) called “the rate of return over cost” (Keynes 1973 [1936], p. 140).

⁸³ Among the other reasons, this divergence is important because it constitutes for Tobin the principal channel of the monetary policy.

The question of equalization of the two sets of prices in the long run has become the dividing line between the two strands that emerged out of the q-theory. On the one hand, equalization of the two prices has become a foundation for the neoclassical approaches. An alternative approach was suggested by Minsky in his two-price theory of investment. As in Keynes, for Minsky (2008 [1986], p. 160, 195-205) one and the same element of the means of production can be treated both as an investment good, i.e. a part of the current output, and as a capital asset having the capacity to yield future profits. This results in two distinct principles of determination of value of these means of production. Their value as an investment good depends upon money wage rates, the productivity of labor, and the markups on the labor costs, whereas the value of the same means of production as capital assets is determined by the cash flows the assets are expected to generate and the capitalization rate. Minsky (1994, p. 25) summarized this insight saying that “a capitalist economy has both a “CPI” and a “Dow Jones””. Unlike Brainard and Tobin (1968) and Tobin (1969), for Minsky (2008 [1986], p. 200-201, 1993, p. 16-17) there is no need for the market value and the replacement costs to equalize, even in the long run. This is due to their being dependent on different factors and determined in different markets. Another departure from Brainard and Tobin (1968) and Tobin (1969) is an explicit account of, first, the borrower’s and lender’s risk and, second, the financing conditions of investment.

In spite of the differences among the theories of investment suggested by Keynes, Tobin, and Minsky, there is a fundamental unifying principle making these theories similar at their core. The three approaches hinge on a distinct conceptualization of capital assets as having two sets of prices – the market value based on capitalization of future

yields and the replacement costs. Moreover, the three theories argue that it is the gap between the two sets of prices that constitutes an inducement to invest.

The debate around the empirical and theoretical validity of these theories of investment offers good reasons to believe that the approaches to investment developed by Keynes, Tobin, and Minsky have a limited explanatory power.

For one, the theories contradict a number of empirically observable phenomena. Comparative econometric and case studies find mixed or weak support for the q-theory of investment (e.g., Fisher, Merton 1984, Bosworth 1981) and show it to be inferior to other investment theories (Clark 1979, Kopke 1982). Crotty argues the inability of the theories of investment developed by Keynes, Tobin, and Minsky to stand empirical tests is not surprising due to the conceptual problems at the heart of these theories. This idea found its sharpest expression in the conclusion that “the financial theories of investment espoused by Tobin, Minsky, and the Keynes of Chapter 12 are simply wrong” (Crotty 1990, p. 538). Although Crotty underestimates the validity of these theories under particular circumstances and thus arrives at a too strong conclusion, he nevertheless puts a hand on an important conceptual mistake preventing these theories from being a general theory of investment. The mistake is the conflation of managers and shareholders, as in the case of Tobin and Minsky, or the domination of owners over managers, as in Keynes. For Crotty (1990, p. 535-537), by contrast, investment decisions can be relatively independent of q. There is a number of reasons why managers and shareholders would have a different approach to investment. First, managers and shareholders respond differently to competition. Facing a decline in profitability due to intensified competition, shareholders would sell the equities to invest in a more profitable enterprise, whereas

managers might invest and build up excess capacity. Second, due to liquidity of investment in securities and irreversibility of investment in real capital, owners and managers have a different tolerance of risk associated with investment projects.

The discussion around the theories of investment developed by Keynes, Tobin, and Minsky offers good reasons to believe that the gap between the market value of the capital assets and the replacement costs cannot be a foundation of a general theory of investment. Does it make the idea of the two sets of prices of capital goods irrelevant?

5.3. From the Dual Price System to the Nature of Financial Profits: Hilferding

The problematic character of the theories of investment developed by Keynes, Tobin, and Minsky lies, however, not in the distinct conceptualization of capital goods and the two sets of prices associated with them, but in an attempt to use this conceptualization to develop a general theory of investment. The relevance of the dual price system goes beyond the theory of investment. A look in the history of thought can offer insights into other possible questions that can be addressed with an aid of the dual price system.

One such set of questions is the theory of capital and capital controversies. Hicks (1974) argued that a possible way of looking at the history of capital controversies – from Ricardo and Malthus to the Cambridge capital controversy – is to understand it through the lens of evolution of two distinct approaches to aggregation of the real capital for the economy as a whole. These two approaches are called by Hicks “fundist” and “materialist”.⁸⁴ For “materialists”, capital is a stock of physical goods. The “volume of

⁸⁴ Hicks considers Smith, British classical economists, Marx, Jevons, Boehm-Bawerk, Taussig, and Hayek to be “fundists”, and Cannan, Walras, Marshall, Pigou, and J.B.Clark – “materialists”.

capital” is aggregated using prices of capital goods themselves. By contrast, for “fundists”, capital is a capacity of producing goods in the future which is embodied in these physical goods. The “value of capital” can thus be measured as capitalized values of future net products. Thus, for Hicks (1974, p. 315), the “fundist” concept of capital is forward-looking, whereas the “materialist” concept is backward-looking. Viewed from this angle, the relevance of Keynes’s two prices of investment is in that it explicitly acknowledges that the concept of capital is simultaneously forward- and backward-looking, making him both a “fundist” and a “materialist”.⁸⁵ Capital is both a sum of values necessary for reproduction of the given stock of capital and a discounted future flow of value accruing to this stock. Keynes’s two sets of prices suggest that the two approaches to capital, which represent two opposing camps in some of the capital controversies, are in fact not mutually exclusive.

Without denying the continuous co-existence of the two approaches to capital, Mehrling (2001, p. 53) shifts the emphasis from continuity to change in the prevalent approach to capital in the 20th century. The forward-looking concept of capital rose to dominance shifting the backward-looking concepts to the margins. In Mehrling’s words, “the economics view” of capital has given way to “the finance view”. In this context Mehrling (2006, p. 73) finds Irving Fisher interesting in that he was “caught with one foot in the 19th century and the second in the 20th”. By this Mehrling means Fisher inherited the classical political economy view, according to which “the accumulation of

⁸⁵ Keynes probably took this idea from Irving Fisher. Hicks (1974, p. 309) mistakenly treats Keynes as mostly a “materialist” with “no more than slight signs” of “fundism”. Hicks clearly underestimated the essential role Keynes attributed to investment as capital assets yielding future revenues, and on these grounds treated only Irving Fisher as simultaneously a “materialist” and a “fundist”. Similarly, Hicks was one-sided treating Marx as a “fundist”, whereas Marx was in fact both a “materialist” and a “fundist”.

capital-wealth from the past determines the flow of future income-services”, but also anticipated the development of the modern theories of finance (CAPM) and the modern macroeconomic theory (DSGE models), according to which “future income flows determine current capital valuations”. An early appreciation of the two modes of valuation of capital stressed by Mehrling is not unique to Fisher. It was anticipated by Marx (1977 [1894]) and Veblen (1904).

For Marx, the shift in the prevailing concept of capital to what he called fictitious capital comes from the emergence of the joint stock companies leading to the titles of ownership becoming commodities. The market value of these titles of ownership is determined differently from the value of the constant and variable capital these titles represent, therefore, there is no economic reason for the two to coincide. Marx (1977 [1894], p. 320) stressed that these titles of ownership

“become commodities, whose price has its own characteristic movements and is established in its own way. Their market-value is determined differently from their nominal value, without any change in the value (even though the expansion may change) of the actual capital. On the one hand, their market-value fluctuates with the amount and reliability of the proceeds to which they afford legal title... But assuming the expansion of the actual capital as constant, or where no capital exists, as in the case of state debts, the annual income to be fixed by law and otherwise sufficiently secured, the price of these securities rises and falls inversely as the rate of interest”.

Thus, anticipating Fisher, Marx stressed the co-existence of the two systems of valuation of capital.⁸⁶ The market value of capital (also called by Marx “fictitious capital”) is determined by discounting the expected flow of surplus value by the ongoing interest rate, while the nominal value of the capital invested is determined by its

⁸⁶ Unlike Keynes, Tobin, and Minsky who were concerned with valuation of the elements of the capital assets (or constant capital, in Marx’s language), Marx focused on the total sum of values invested, i.e. a sum of the constant and variable capital.

reproduction costs, or the socially necessary labor time necessary to reproduce the elements of the constant and variable capital.⁸⁷

The significance of the transformation of the prevalent concept of capital from a backward- to a forward-looking stressed by Mehrling goes beyond the mere question of capital valuation. This transformation has also brought about a new form of profit making. An emergence of a new form of profit is what makes this transformation of the concept of capital socially relevant. An observation that the gap between the two sets of prices created a scope for extraction of a particular type of profit is Hilferding's theoretical innovation that has not received its due appreciation.

Hilferding explicitly associated the gap between the two sets of prices with a particular type of profit, namely, founder's profit. Being concerned with a genesis of the joint stock capital, Hilferding considered money capital K invested in purchasing the means of production and labor power. If the average rate of profit is r , the expected flow of surplus value generated per period in the process of production would be equal to $\pi = r \cdot K$. Following Marx, Hilferding assumed that the rate of profit systematically exceeds the interest rate, $r > i$, due to a different social status of money lenders and active capitalists. Given that from the perspective of an owner of money capital buying shares is akin to lending, Hilferding argued that shareholders would be satisfied with the rate of interest, as opposed to the rate of profit, and that they would be willing to pay $\frac{\pi}{i}$ for the

⁸⁷ For Marx, the nominal value of capital is determined by the present conditions of its reproduction, not its historical costs, and is therefore not backward-looking, as in Hicks, but "present-looking". The same emphasis can later be found in Tobin and Minsky.

shares.⁸⁸ From the rate of interest being lower than the rate of profit ($r > i$), it immediately follows that discounting the future stream of surplus value with the former would necessarily result in the market value of an incorporated enterprise standing above the value of the actual capital invested ($\frac{\pi}{i} > \frac{\pi}{r}$). The difference between the market value paid by investors and the initial capital invested, $\frac{\pi}{i} - \frac{\pi}{r} = \frac{\pi}{i} - K$, constitutes founder's profit.

A relevant observation for our purposes is that founder's profit thus understood represents the difference between the two sets of prices of capital – between the market price of the capital asset and the cost of reproduction. Even if the dual price system does not offer a framework for a general theory of investment, the conceptualization of capital goods as having two sets of prices is relevant, because it sheds the light on the nature of a peculiar form of profit.

5.4. Reading Minsky through Hilferding's Lens. What are Capital Gains?

The shift of focus to the character of profit offers a useful lens for revisiting Keynes, Tobin, and Minsky. Although none of them had a theory of profit associated with the gap between the two sets of prices, Minsky has insights that can further our understanding of financial profits.

Minsky (1986) developed his argument in the context of profits from mergers and acquisitions. In spite of this specific focus, the general thrust of his argument is clearly

⁸⁸ Hilferding also considered risk premium that would bid the rate of return on shares above the interest rate. However, he assumed that competition would eliminate the risk premium. On these grounds he dropped the risk premium in his ultimate analysis of founder's profit.

applicable to a broader range of profits associated with functioning of the capital markets, e.g., capital gains and underwriting revenues. He opened up the discussion with Keynes's (1973 [1936], p. 158) distinction between speculation and enterprise. Minsky maps this distinction onto that between the corresponding types of profit – something not actually argued by Keynes himself. Profit from speculation is taken by Minsky to arise due to appreciation of the value of the assets, profit from enterprise comes from the income the assets earn in production. Minsky (1986, p. 349) argued that “Keynes's sharp distinction between speculation and enterprise is not wholly warranted”. For Minsky, any enterprise is motivated by capital gains, because there is an implicit capital gain realized at the moment of investment, whenever the price of a capital asset exceeds the cost of investment (Minsky 2008 [1986], p. 238-239).⁸⁹ To explain this gain, Minsky (1993, p. 13) argued that at the moment of purchase of an investment output there is a change in its value – from its sale price to capitalized earnings it is expected to produce. This results in a capital gain. On these grounds, he shifted the focus away from the distinction between speculation and enterprise. The relevant fact for Minsky (1986, p. 348-351) is that in the case of mergers and acquisitions financial markets transform future increases in market power (or in labor productivity, as one might want to add) into immediate capital gains. This is in line with his general view of financial markets as mechanisms transforming expected future cash flows associated with capital assets into a set of current prices (Minsky 2008 [1986], p. 194).

⁸⁹ Minsky disagrees with a point Keynes himself never made. Keynes did not try to map his distinction between speculation and enterprise onto the corresponding types of profit. Keynes (1973 [1936], p. 151) was aware that investment can be driven by “an immediate profit” from floating it on the stock exchange, as opposed to a series of future flows of value.

Thus, for both Hilferding and Minsky, at the most abstract level the gap between the market price and the reproduction costs of an enterprise represents a particular form of profit, namely, founder's profit and gain from mergers and acquisitions, respectively. Hilferding and Minsky agree on several characteristics of capital gains. First, an analysis of capital gains should not start with a presupposition that they are speculative. Second, capital gains should rather be linked to creation of surplus value (for Hilferding), or to the expected future yield of the capital assets (for Minsky). Thus, third, functioning of the capital markets allows one to realize gains associated with value that is not yet created.

In spite of these similarities between Hilferding and Minsky, there is a fundamental difference between their approaches. Specifically, they offer different explanations of the existence and persistence of the gap between the two sets of prices.

For Minsky (2008 [1986], p. 200-201, 1993, p. 16-17), the two sets of prices, although related, do not have to be equal due to their being dependent on different factors and determined in different markets. Furthermore, while the price of current output is in principle sluggish, the price of capital assets is inherently volatile (Minsky 1975, p. 92). Therefore, the two sets of prices can and do vary independently of each other, thus, resulting in endogenously generated cycles.

Hilferding's position combined the arguments by both Marx and Veblen. Recall that for Hilferding it is precisely the difference in the rates of return required by shareholders and active capitalists that creates a systematic divergence between the market capitalization and the actual capital invested. Hilferding followed Marx

emphasizing the difference in the rates of return.⁹⁰ At the same time, in the subsequent analysis Hilferding (1910) also emphasized the difference between the information possessed by managers and shareholders, resulting in a different valuation of the same enterprise. In this respect he is close to Veblen (1904, p. 77) who stressed a discrepancy between the actual and the putative earning capacity of the enterprise. Thus, the problem for Veblen boiled down to the difference between managers and investors – insiders and outsiders. Given that the latter do not know the actual earning capacity and that the former can, furthermore, manipulate or withheld relevant information, it is not surprising that the valuation of the company by investors would exceed that by managers. This discrepancy between the actual and the putative earning capacity resulted in the value of the business capital exceeding the actual capital invested in the process of production by the amount of good-will and other intangible assets (Veblen 1904, p. 60, 72-73).

Thus, Minsky differs from Hilferding by not connecting the gap between the two sets of prices with a difference in the rates of return. Moreover, there are reasons to believe Minsky would have been critical of this explanation of capital gains, because he criticized Keynes on similar grounds. Keynes (1973 [1936], p. 137) argued that investment would be carried out until marginal efficiency of capital is equal to the rate of interest, or until the supply and demand price of investment are equalized. He considered these two approaches to be equivalent. Minsky (1975) was critical of seeing them as equivalent and he favored the approach based on a comparison of the two sets of prices of

⁹⁰ Hilferding also echoes Marx's (1867 [1844], p. 477) conclusion that the value of the imputed capital-values "may fall or rise quite independently of the movement of value of the real capital for which they are titles".

the capital goods. For Minsky, the difference in the rates of return was an evidence of Keynes's being trapped in the classical economics.

Hilferding's argument, however, shows that the emphasis on the differences in the rates of return does not need to be associated with the neoclassical economics. In fact, the difference in the rates of return is an important aspect of Marx's theory, because it corresponds to a different social status of lenders and active capitalists. The significance of the difference in the rates of return is also appreciated by Tobin (1969). His q-theory of investment is usually associated with equalization of the rates of return, but this reading of Tobin overlooks the importance he assigned to the divergence of the rates of return. He was very clear that the divergence between the rate of profit and the rate of interest is the principle channel of the monetary policy that operates through revaluing the market value of capital in relationship to its reproduction costs. Thus, for Tobin, without the gap between the rates of return monetary policy would not have been possible.

The difference in understanding the reasons of existence of the gap between the two sets of prices has an important implication.⁹¹ Given that Minsky does not have a convincing explanation why the gap would not be eliminated in the long run, in his approach financial markets remain a black box creating capital gains by a mere act of changing the mode of valuation of the capital goods. There is no social content to it. Nor is there a discussion of the macroeconomic sources of the gain associated with the change

⁹¹ In spite of the differences, Hilferding and Minsky agreed that there is a deeper conceptual distinction between the two sets of prices. Unlike the price of produced commodities, the price of capital assets is "the price of a revenue" (Hilferding 1910, chap. 7) or "prices for future streams of incomes" (Minsky 1993, p. 10). Thus, the price of capital as assets is actually not a price of capital, but rather of the flows of income associated with this capital.

in valuation.⁹² By contrast, focusing on a different social status of active capitalists and shareholders, Hilferding explains the origin of the gap between the two sets of prices, which in turn allows him to uncover the social relations allowing for this change in the mode of valuation. Thus, what passes for a solution to the problem of the nature of capital gains for Minsky – the transformative nature of the capital markets – itself requires an explanation for Hilferding.

There is, however, a deeper reason that can possibly explain why the capital markets can remain a black box for Minsky, whereas Hilferding finds it necessary to focus on the social relations behind it. A possible root of the difference between the two approaches lies in the theories of value supported by Minsky and Hilferding. Whereas the existence of capital gains does not challenge Minsky's theory of value, capital gains do pose a major challenge for Hilferding's theory of value. This could be a reason why Minsky could be satisfied with the transformative function of the capital markets, whereas Hilferding needed to explain this function itself as a change in the social relations. Consider the difference in the theories of value in a greater detail.

For Minsky, an ability of the capital markets to change the mode of valuation of the capital assets results in capital gains being indistinguishable from profits from production. But this is not a problem for him, especially in his early work (Minsky 1975) in which he inherited Keynes's theory of capital and value, according to which profit from production itself appears as nothing more than a capital gain.

⁹² By not providing a reason why the gap would be systematically maintained, Minsky welcomes explanations according to which the two sets of prices would equalize, at least in the long run. This is impossible in Marx and Hilferding, for whom the gap would persist.

Keynes (1973 [1936], p. 135, 213-214) analyzed profit from production as quasi-rents, or yield on capital assets – the difference between the total returns from selling the output produced with the aid of the capital assets and the technically determined costs of production.⁹³ The quasi-rents exist due to scarcity of the capital assets. Given that this scarcity is determined outside of the process of production in which these capital assets are deployed, there is no immediate connection between the magnitude of the quasi-rents and the costs of production. Therefore, profit from production itself appears in Keynes as an abstract return on capital assets, i.e. as a form of a capital gain. This is reminiscent of the labor commanded theory of value advanced by Adam Smith.⁹⁴ According to Smith, the value of the output is not the value embodied in the commodities produced, it is rather the value this output can command, i.e. be exchanged for. Skidelsky (1992, p. 326) made an even stronger argument claiming that in *The Treatise on Money* Keynes “had no real theory of profit”, i.e. no theory of what determines “the rate of return on physical capital”. This observation holds for *The General Theory*, as well. It was noticed by Joan Robinson (1985, p. 158) when she argued “the formation of an overall rate of profit is left hazy”.

If profit from production is studied as akin to capital gains, it is not surprising that for Minsky capital gains are reminiscent of profit from production and that one cannot distinguish between profits from enterprise and profits from speculation. In this sense, Keynes’s theory of value and capital also adopted by early Minsky made this conclusion

⁹³ Minsky (2008 [1986], p. 228) suggested Marshall and Keynes used the concept of quasi-rents to draw a parallel between the returns on capital assets in the sphere of production and ground rent in Ricardo.

⁹⁴ I am thankful to Tomas Rotta who drew my attention to the similarity between Keynes’s theory of value and Smith’s labor commanded theory of value.

inevitable. An important point here is that the distinction between profit from production and from functioning of the capital markets vanishes not because there is no difference between them, as suggested by Minsky. It is rather because Keynes, and then following him Minsky, mistake profit from production for a sort of capital gain. It is particularly clear in Minsky's take on a firm as a cash-flow machine (Minsky 2008 [1986], p. 228), and his argument about investment in capital assets being akin to buying a bond (Minsky 2008 [1986], p. 238). This view culminates in Minsky's (1975, p. 132) statement that "in Keynes's asset-valuation model, productive capital assets are perhaps best viewed as another, albeit peculiar, speculative financial asset".

If there is no distinction between profit from production and capital gains, and both arise in the sphere of exchange, treatment of the capital market as a transformative device does not pose a major problem for a theory. There is no inconsistency between profits from production arising in the sphere of circulation and capital gains arising through a particular mechanism of the sphere of circulation, namely, the capital market. Thus, the capital market can indeed remain a black box.

This approach stands in sharp contrast with Hilferding's. Following Marx, he considered profit to be a form of surplus value arising through exploitation. If profit is understood as a form of subdivision of the flows of the currently produced value, founder's profit poses a theoretical problem. What is the connection between value and founder's profit that does not seem to bear any relationship to exploitation in the sphere of production? Thus, the labor theory of value forces Hilferding to pose a question of the social relations in order to reconcile founder's profit with the flows of value.

Hilferding (1910, chap. 7) solves this theoretical conundrum by treating founder's profit as an economic category sui generis associated with a transformation of capital. At the most abstract level, the genesis of the corporate form of the business organization involves a transformation of the profit-bearing capital into interest-bearing capital. Another way of looking at this process is as a transformation of the nature of capital from individual capital into truly social. In the 19th century an individual capitalist invested in the process of production and retained control over it. With a spread of the corporate form of business organization ownership of capital becomes diffused across the society, and the amount of capital invested depends on the temporarily idle funds scattered across the society as opposed to funds of an individual capitalist. It is the social nature of the joint-stock capital that makes possible to turn certain assets into assets with a greater value.

For Hilferding, a counterpart of this change in the nature of capital is a systematic divergence between the rate of profit established in the sphere of production and the rate of return required by shareholders. Without the difference in the rates of return, the capital markets would not be able to transform the future gains of production into the current capital gain. The conclusion that capital gains cannot exist without the difference in the rates of return becomes apparent if capital gains are presented

as $KG = \frac{r \cdot K}{i} - K = \left(\frac{r}{i} - 1 \right) \cdot K$, where KG is a capital gain (founder's profit), r – the rate

of profit, i – the rate of interest, K – sum of value. From this it immediately follows that for capital gains to be positive, the rate of profit should stand above the rate of interest. If the rates of return are equal, there would be no capital gains, and the asset seller would

only recover what had been invested.⁹⁵ Thus, even though Hilferding makes an argument similar to Minsky suggesting that founder's profit is a lump sum gain equal to the discounted expected profit of enterprise, Hilferding does not leave the explanation of how exactly that happens to the black box of the capital markets.

To recap, if one adopts a theory of value in which profits from production are theorized as some sort of a capital gain, as in Keynes and Minsky, the difference between profit from production and capital gains disappears. The description of the capital markets as a black box does not stand in conflict with the rest of the theory. From the standpoint of the labor theory of value, however, profit from production is understood as a form of surplus value arising due to exploitation of workers, as in Hilferding. To explain the possibility of capital gains, the labor theory of value forces one to analyze the social relations making capital gains possible. The strength of Hilferding's approach is that he simultaneously retains the difference between profit from production and capital

⁹⁵ Capital gains understood in this way are a particular case of a broader category of profits that do not have a counterpart in the currently produced output. There are two types of profit involving re-division of monetary assets. The difference between these types of profit hinges on the conceptual distinction between lending commodities as use-values and lending commodities as values, introduced by Teixeira and Rotta (2010). When a commodity is loaned as a sum of values, it generates capital gains of the type discussed here. These capital gains hinge on the difference between the rate of return on the sum of values and the rate of return required by the holder of the claim. By contrast, when a commodity is loaned as a use-value, the concept of a rate of return on its value becomes meaningless. Therefore, there is also no role to be played by the difference in the rates of return, and the sum total of profit from dealing in these commodities is reduced to a "pure form" – a simple difference between the current market price determined as a rent per period discounted with a required rate of return and the reproduction costs of this commodity (in this case, equal to zero). Thus, the resulting profit is reduced to $KG = \frac{Q}{i} - 0 = \frac{Q}{i}$, where Q is the expected future rent per period, i is the ongoing interest rate. This pure form of capital gains requires a separate study and is not addressed in the present article. The paradigmatic case of lending commodities as use-values is given by valueless commodities, such as land and knowledge commodities. For a compelling analysis of valueless commodities, see Teixeira and Rotta (2010). The issues involved in determination of value of non-commodity money characteristic of modern capitalism raises a question whether capital gains from dealing in money should be theorized as profit of type one or type two above. To address this question, one would first need to establish a relationship between the value of money and the value of credit as an asset whose exchange value is determined by the current value of the future money payments it promises to make (Mehrling 1996, p. 332). Complications for the nature of capital gains that arise with valueless money lie beyond the scope of the present work.

gains, yet also shows the connection between the two. The connection is important, because it implies that founder's profit remains a specifically capitalist form of profit. The very act of discounting of the future flow of surplus value hinges on an expectation that flows of surplus value will be systematically created. Therefore, founder's profit bears an important, though mediated relationship to surplus value creation. At the same time, Hilferding argues that founder's profit cannot be conflated with surplus value. Thus, the labor theory of value forces Hilferding to open the black box of the capital markets and examine the relations involved in the transformation of the future surplus value into the current founder's profit.

5.5. Characteristics of Capital Gains: Beyond Hilferding and Minsky

The discussion of capital gains by Hilferding and Minsky sheds the light on some characteristics of financial profits.

First, a gap between two sets of prices of capital assets – their price as an interest-bearing asset and their price as a value-containing commodity – constitutes a space for capital gains. Nevertheless, capital gains should be understood not simply as a result of a transformative function of the capital markets, as in Minsky, but rather as an outcome of a particular set of social relations, as in Hilferding. Thus, Hilferding relates to Minsky in the same way as Marx relates to Smith. In particular, Marx showed that, contrary to Adam Smith, markets and private property alone cannot create a capitalist mode of production. Similarly, the existence of the capital markets alone cannot explain the nature and mechanism of extraction of capital gains.

Therefore, second, the existence of the type of capital gains considered here presupposes a set of specifically capitalist social relations.⁹⁶ Capital gains develop as a systematic and a large-scale phenomenon on the basis of capitalism for a number of reasons. Capital gains hinge on discounting future flows of surplus value produced with an aid of the capital asset. Thus, they presuppose a continuous production of surplus value. Furthermore, the possibility of discounting future returns with an interest rate presupposes systematic borrowing and lending and development of the money market. These in turn hinge on systematic leakages of value from the circuit of capital and revenue, necessitated by the logic of the circuits themselves.⁹⁷ Finally, capital gains require the interest rate to stand below the rate of profit, which is also an outcome of a prevalence of the capitalist mode of production, as opposed to usurious interest rates characteristic of the earlier stages of historical development. Thus, behind markets as the source of capital gains are the differences in the required rates of return.

Not only does the existence of capital gains presuppose a certain level of development of the capitalist relations, but also the emergence of the two sets of prices itself hinges on a specifically capitalist process of converting use-values into commodities. Rotta and Teixeira (2011) show how the commodity form attaches to the human capacity to create value, to land, and to knowledge, resulting in, respectively, commodities labor power, land, and knowledge-commodities. In the case of capital gains, commodity form attaches to capital itself.

⁹⁶ Capital gains can exist in non-capitalist systems, as long as there are market relations.

⁹⁷ This process of hoard formation analyzed by Marx (1967 [1885], p. 158-159, 163-166) is further developed by Hilferding (1910, p. 67-81), Itoh (1988, p. 259-260, 401), and more formally by Lapavistas (2000).

From a Marxist perspective, there is nothing surprising in the co-existence of two prices of capital, and the very possibility of such a divergence is contained in Marx's distinction between use value and exchange value. Marx himself uses this insight in order to explain the origin of surplus value. Commodity labor power has a use-value (ability to create value) and value (costs of reproduction of the commodity labor power), and it is this difference that explains the origin and persistence of surplus value. Similarly, capital has a use value (its ability to extract surplus value) and exchange value (reproduction costs of the elements of constant and variable capital), and it is the gap between the two – between the discounted value accruing to the owner of the capital asset and the reproduction cost of the capital asset – that creates capital gains. Thus, there is an affinity between valuation of labor and capital, between surplus value and capital gains.⁹⁸ In the final analysis, both forms of profit are made possible by the distinction between use value and value.

Fourth, capital gains do not have a counterpart in the current output. Precisely because the financial markets transform the future gains into the present capital gains, as stressed by Veblen, Hilferding, and Minsky, there is no real value created at the moment of this transformation. This is why, contrary to Minsky (1993, p. 13), it is misleading to view a simple act of purchasing an investment good as being capable of creating value.⁹⁹

⁹⁸ There is a difference between the two cases of labor and capital due to the role of time. For the labor power, its ability to create value and its reproduction cost are simultaneous and both occur in the current period. By contrast, while the reproduction cost of capital is also current, the discounted flow of surplus value is a future flow.

⁹⁹ In the case of capital gains narrowly defined, the absence of a counterpart in the current output is acknowledged by the system of national accounts which excludes capital gains from calculations of GDP.

Capital gain is a form of advance, or prepayment, and in this sense the profit arising through the differences in the rates of return can be thought of as ‘fictitious profit’.

The lack of a counterpart in the current output, however, does not prevent these profits from being monetized, which poses a question of the macroeconomic source of capital gains. As was shown in Lapavitsas and Levina (2011), at the moment of accrual capital gains come from re-division of stocks of money of the asset buyers.¹⁰⁰ In this sense, capital gains are a form of profit upon alienation (Steuart) characteristic to the pre-capitalist stages of historical development. Therefore, the criticisms of capital gains as somehow not real overlook the point that capital gains involve redistribution of stocks of money, thus, have very real redistributive outcomes. Capital gains provide their recipients with purchasing power indistinguishable from income having a counterpart in the current output.

To recapitulate, capital gains take the form of a gap between the two sets of prices of capital assets. At their core, capital gains systematically arise only on a capitalist foundation, but simultaneously they involve predatory relations with respect to the ultimate asset holders. Thus, paradoxically, the most developed form of surplus extraction in capitalism is simultaneously a reinstatement of an antideluvian form of profit-making prevalent in earlier modes of production.

¹⁰⁰ This was the key point of criticism of Hilferding by Itoh (1988, p. 287). Itoh argued that founder’s profit should be understood as coming from re-division of monetary assets of a security buyer, not as a part of the surplus value accruing in a lump sum.

5.6. Capital Gains Come in Many Forms

Capital gains as a difference between the market price and the costs of reproduction of capital goods represent the most abstract and general category of apparent wealth creation. Founder's profit is a logical origin of a broader range of financial profits, including profit from mergers and acquisitions.¹⁰¹ This form of profit attaches to a range of marketable objects having expected future cash flows associated with them.

It is also a pure form of a broader range of profits, because there is no service that can be associated with extraction of capital gains. Capital gains, however, come in many forms. Underwriting revenues, fees from mergers and acquisitions, securitization revenue, trading gains (both proprietary and on behalf of clients), fees from asset management, and even managerial bonuses in the form of stock options are other particular forms taken by capital gains, broadly defined. The mainstream theories and, hence, national accounting focus on provision of useful services associated with these activities and on these grounds treat them as value creating. The approach developed above suggests that these activities, although indeed involving a useful service, do not create value, at least not at the moment of their accrual. These revenues rather come from redistribution of monetary assets, facilitated by the differences in the required rates of return. Therefore, these seemingly different profits are capital gain-like. Monetary assets are used here as a generic category for monetary wealth that takes the form of either the

¹⁰¹ Minsky (1986) was aware that that the market valuation of equities constitutes a starting point for an analysis of mergers and acquisitions, and in this sense he can be argued to generalize Hilferding's concept of founder's profit onto other forms of financial profit.

already existing stock of money or the newly created credit. The latter happens when an asset is purchased using borrowed funds.

The multiplicity of forms of capital gains has an important implication. The shift of focus from investment to the nature of capital and capital gains helps clarify issues of validity of Keynes-Tobin-Minsky theory of investment. The dual price system has a broader relevance than just its connection with investment, because investment is not the only mechanism for extracting capital gain-like profits. Furthermore, a theory of capital gains shows conditions under which there would be a connection between the two price system and investment, as suggested Keynes, Tobin, and Minsky, and under which their theories are misleading, as argued by their critics. Their approaches are valid when an agent making an investment decision is remunerated by a capital gain. This is the case with a founder setting up an enterprise to float it off (a venture enterprise) and to some extent with managers remunerated by stock options. When an agent making an investment decision is not remunerated by an associated capital gain, the explanatory power of Keynes-Tobin-Minsky theories is limited, because under those circumstances an investment decision is a much more complex process than comparing the market price and the reproduction costs of the capital assets.

5.7. Forms of Profit due to the Gap between the Two Sets of Prices of Capital and the Transformation of Banking

The gap between the two sets of prices of capital goods was shown to be a source of a particular form of financial profit. The remainder of this chapter will focus on the concrete forms this profit takes and on the mechanisms of its extraction, which will in

turn be used to reevaluate the transformation of banking. It will be argued that the significance of the transformation of banking cannot be fully recognized on the basis of bank functions. Some changes in banking are associated with a much more profound transformation of the character of bank revenues than functions. The logical origin of this argument can be established by examining the turn from commercial banking to investment banking featuring a paradox that reappears in other banking activities.

5.7.1. Profit from IPO

The simplest form of extraction of capital gains is an initial public offering (IPO). An IPO is a process of issuing equity shares representing claims on future revenues and selling them to investors. A decision to go public can be made for different reasons, and a specific combination of factors would vary from case to case. The most common reasons include owner's decision to cash out, debt repayment, covering operating expenses, capital expenditures, and augmenting corporate treasury. The usual debate on the reasons of existence of IPOs is nevertheless typically centered on two main approaches.

According to the first approach, an IPO is a means of raising funds necessary for the business expansion. In some cases IPOs were used for this purpose. For example, in the 1990s IPOs were an important source of finance for start-ups in the high-tech sector. Nevertheless, outside of these cases, IPOs were more often than not used by companies that had sufficient cash flows to finance their operations.¹⁰² Moreover, even when there is

¹⁰² For the most recent example, see a debate around Facebook going public. John Gapper (2012), Facebook Ought to Ditch its Public Offering, Financial Times, February 2, 2012, available at <http://www.ft.com/intl/cms/s/0/cd03c402-4dba-11e1-a66e-00144feabdc0.html#axzz11SKpGLxa>. Sarah Gordon (2012), The Numbers are Vast, but Beware the Bite of Reality, Financial Times, February 3, 2012, available at <http://www.ft.com/intl/cms/s/0/aebe570c-4db3-11e1-b96c-00144feabdc0.html#axzz11SKpGLxa>. For an insightful discussion, see Mason, JW (2012), The Capitalist

a need to raise funds, equity issuance is one of the most expensive ways of doing it, not only because of the transaction costs, but also because of the adverse signals to potential shareholders suggesting a lack of confidence in the ability to meet debt-service payments through future earnings. For these reasons, raising funds cannot be the key reason of existence of IPOs.

An alternative approach explains IPOs by the initial investors' and stakeholders' desire to cash in, that is, to substitute their sunk capital for money, the most liquid form of wealth. Navin and Sears (1955, p. 127) consider this strive for liquidity to be "the trump card" beating other reasons behind the historical origin of IPOs. They argue "the years 1887-1902 produced a solution to one of the troublesome problems created by the industrial revolution, the problem of capital inflexibility in the industrial segment of our economy... By seeking incorporation the proprietors of industrial enterprises had acquired, through the issuance of common stocks, a potentially easy means of transfer" (Navin, Sears 1955, p. 136). The strive to liquidity and associated separation of ownership and control are sometimes treated as the reason behind the entire history of IPOs, not just their origin. For example, Lazonick (1992, p. 451, see also 2007, p. 1021, 2004, p. 19) emphasizes that "contrary to shareholder folklore", "common share issues have never been important in US industry as a means for financing enterprise expansion". The money was rather normally used "to permit owner-entrepreneurs to take their leave of the firms that they had built up without disrupting the continuity of the enterprises" (Lazonick 1992, p. 450-451). Thus, in the light of the actual historical experience a flight to liquidity is a more universal aspect of IPOs than raising funds. This idea is sometimes

Wants an Exit, Facebook Edition, available at <http://slackwire.blogspot.com/2012/02/capitalist-wants-exit-facebook-edition.html>.

put abstractly, namely, that capital strives to exist in the most liquid form, so that an IPO is a concrete mechanism of accomplishing this goal by releasing capital from a connection to a particular circuit of capital. This argument plays an important role in Marxist literature. Uno school of Marxism views development of forms of capital as an attempt of capital to set itself free from limitations of particular use-values (Itoh 1988). Kotz makes a similar argument but not at an abstract theoretical level, as in Itoh, but in the context of a rise of bank control of the U.S. corporations since the late 19th century. For Kotz (1978, p. 148-149), a rise of the financial control of corporations by banks implies a shift of ultimate power in capitalism from individual capitalists who invested in the process of production, owned, and controlled it to banks that own and control “abstract capital”. Banks, unlike industrial capital, are not tied to any particular industry and rely on other people’s capital as a principle of their operation. With their becoming a center of capitalist control, capital seeks to escape the confines of particular use-value production through taking on a more general form. A similar argument from a broader historical perspective can be found in Braudel (1983, p. 113, 232, 372-373) who treats the sphere of circulation as the natural terrain for capital which originated in the sphere of commerce, and the stock market as a mechanism of reinstating convertibility of illiquid forms of capital into liquid ones.

A useful way to think about these two approaches to IPO – as a mechanism of raising funds and of obtaining liquidity – is through the lens of the concepts of funding liquidity and market liquidity. These notions developed in the context of functioning of security dealers became well-known through the work of Brunnermeier and Pedersen (2009). Even though widely deployed, these concepts haven’t acquired a universally

accepted meaning, and different authors mean different yet related things. For example, for Borio (2010, p. 71), similar to Brunnermeier and Pedersen (2009), funding liquidity seems to be a broader category than market liquidity, with the latter being one of the conditions for the former. In particular, for Borio, market liquidity is “the ability to trade an asset or financial instrument at short notice with little impact on its price”, whereas funding liquidity is “the ability to raise cash (or cash equivalents) either via the sale of an asset or by borrowing”. The same approach is used by CGFS (2011). Mehrling (2011, p. 26) treats market liquidity as the capital market liquidity and funding liquidity as the money market liquidity. Tirole (2011, p. 288-290), however, shifts the focus to market liquidity as liquidity of the asset side of the balance sheets, and funding liquidity as liquidity of the liability side. For our purposes it is useful to distinguish between funding liquidity and market liquidity as forms of liquidity demand from the perspective of the size of the balance sheets. Funding liquidity would then involve an increase in liabilities (an expansion of the borrower’s balance sheets), whereas market liquidity – a change in the composition of assets or liabilities, with the sum total of the assets and liabilities remaining the same.¹⁰³ Market liquidity can be accessed through a sale of an asset or a substitution of one form of liabilities by another. This distinction helps to see that the two common approaches to IPOs are about liquidity. When an IPO is explained by a need to raise funds, it effectively means that IPOs are driven by a need for funding liquidity, whereas an approach to IPOs stressing a desire to cash out hinges on a particular form of

¹⁰³ The rise of capital markets provides social foundations of the shiftability view of liquidity suggested by Moulton (1918). Therefore, the triumphs of the shiftability view emphasized by Mehrling (2011) are a counterpart of the development of the capital markets. The shift in the approach to bank liquidity – from the commercial loans view according to which liquidity is assured by cash flows of the ultimate borrower to the shiftability view in which liquidity depends on the state of the capital markets – then becomes just a particular aspect of a broader transformation of the concept of liquidity associated with a rise of the capital markets.

market liquidity – the one substituting active capitalists’ debt to himself (own capital) by equity shares.

From the perspective of the present study the two approaches to IPOs – as a mechanism of raising funds and of obtaining liquidity – have a grain of truth to them, but neither of them goes to the heart of the matter. Given that capitalists’ most general objective is making profit, a useful way to think about IPOs *raison d’être* is profit extraction. This approach nests the two explanations above as its particular aspects.

First, if IPOs are seen as a mechanism for extraction of an immediate gain (founder’s profit), it becomes less surprising that in the majority of cases firms going public have no need to raise funds. At the same time, the desire to extract founder’s profit is not at odds with the need to raise funds, which occasionally accompanies IPOs. In addition to founder’s profit, an IPO followed by business expansion can be a source for profit making in the future.

Second, a conversion of sunk capital into money is a necessary mechanism of founder’s profit extraction, and in this sense the liquidity-based explanation of IPOs supplements the argument based on founder’s profit. Nevertheless, the strength of the approach focusing on profit extraction as a driving force of IPOs lies in its emphasis on liquidity as a means to an end of profit making, not an end on its own.

Incidentally, even the authors emphasizing the drive of capital to acquire a liquid form have hints indicating their appreciation of the profit motive behind it. For example, Navin and Sears (1955, p. 127) argue “the trump card was immediate liquidity at a price the owners probably never imagined their stock to be worth”. The independent promoters’ “ambition was to make quick profits from ripening opportunities” (Navin,

Sears 1955, p. 130). Lazonick (1992, p. 463, 477) also suggests that IPOs in general is a form of rewarding the new venturers. This form of remuneration has acquired a particular significance in the course of “financialization of venture capital” which became more like the money management business (Lazonick 1992, p. 479). For Lazonick (2007, p. 1006), “the prospect of a quick and lucrative IPO or private sale has generated too much of an inducement to venture creation at the ultimate expense of the speculating public. The dot.com boom of the late 1990s was particularly problematic because of the extent to which US households had become active participants in the highly liquid stock markets”.

If founder’s profit is an end of IPOs, what explains the possibility of its systematic extraction? Founder’s profit is often associated with fraud and asymmetric information between insiders and outsiders. This explanation is often identified with Veblen (1904, p. 29-30, 77-80) who put a lot of emphasis on these factors (Wray 2009, p. 812, Bolbol, Lovewell 2001, p. 533-534). Fraudulent practices have always played an important role in IPOs bidding up founder’s profit, since the days of Jay Gould, Daniel Drew, and James Fisk in the late 19th century, when they bought control of a railroad, looted its assets, and sell out before the damage became known by the general public (Kotz 1978, p. 28). Nevertheless, if fraud and misinformation were the only sources of the founder’s profit, how could one explain the fact that founder’s profit existed even when fraud and misinformation were ruled out? Navin and Sears (1955, p. 123, 132-133) argue that in the early days of IPOs and mergers the shares were distributed mainly among capitalists themselves who had skills and knowledge of business. Among the buyers of securities were often commercial distributors of an enterprise going public. Although the chances of fraud and information asymmetry were thus minimized, founder’s profit was

still present. It suggests that differences in expectations about the earning capacity cannot be the only or a key reason behind the market capitalization systematically exceeding the capital invested, hence, nor can fraud and information manipulation.¹⁰⁴ Therefore, fraud and misinformation, as important as they are, cannot be a conceptual starting point of analysis. The origin of founder's profit should be sought elsewhere, after which fraud and misinformation can be brought into an analysis as symptoms of a deeper structural problem and as additional factors further raising founder's profit.¹⁰⁵ In fact, Veblen (1904, p. 62, 68-76) was well aware of this fact, and in his study of founder's profit he singled out factors other than fraud making founder's profit possible. These other factors however were not appreciated enough by many of his followers.

This is when Hilferding's insight becomes relevant. He aims to replace the moral judgment by an economic explanation of founder's profit which for him is "not itself a swindle, although it certainly makes swindles possible" (Hilferding 1910, chap. 7). He treats founder's profit as a specific economic category by focusing on factors that cannot as a rule be prevented by law and rather require a more profound change in the way the economy operates. He argued when an enterprise is incorporated, its mode of valuation changes. The flow of value that is expected to be generated is no longer confronted by an active capitalist whose place is taken by a lender who requires a rate of interest on his investment which stands below the rate of profit. This difference in the rates of return

¹⁰⁴ At the same time, the limits to founder's profit from distributing securities among capitalists offer reasons why the capitalist class in its pursuit to increase this profit would be interested in drawing in general public in purchasing securities.

¹⁰⁵ An emphasis on factors other than fraud and misinformation as the key source of founder's profit is important because it suggests this profit can not be eliminated by tighter legislation, better information dissemination, and fraud prevention measures. This profit arises due to structural characteristics of the modern business organization and therefore elimination or reduction of this profit would require a structural reform.

allows the founder to appropriate a part of the money-lender's loanable capital. This raises a question why would an investor accept a lower rate of return?

One explanation would be the difference in risk borne by an active capitalist and an investor. The risks for a lender are lower than the risks for an active capitalist due to seniority of debt obligations, whereas an active capitalist is a residual claimant on profits produced. This however cannot explain why an investor would require a rate of return below the profit rate, because in fact upon incorporation a shareholder becomes a residual claimant whose risks are identical with risks of an active capitalist.

Another explanation would be liquidity premium. A process of incorporation converts illiquid assets tied up in production into marketable securities. Thus, investors would accept a lower rate of return than individual active capitalists because the assets of the former are more liquid than the assets of the latter. A problem with this argument is that although there is indeed a difference in liquidity which can explain some of the difference in the required rates of return, it does not explain on which grounds it is a founder of an enterprise, a promoter, and an investment banker who appropriate this liquidity premium. Although an IPO converts an illiquid asset into liquid, the liquidity of the asset stems not from the act of IPO itself, but rather from existence of the market in securities and a group of investors. Therefore, even though some of the difference in the rates of return could be explained by the difference in liquidity, there still remains a conflict between the public nature of liquidity and the private appropriation of its benefits.¹⁰⁶

¹⁰⁶ If an IPO were to be handled by a public agency, as opposed to a private profit-seeking institution (an investment bank), founder's profit stemming from a difference in liquidity would be a public gain belonging to all the people.

This leaves us with another explanation of the origin of founder's profit, which seems to offer a better mileage in understanding this form of profit. Shareholders accept a rate of return below the profit rate because they lack control over business enterprise. Control over business enterprise is therefore the key difference between the social status of shareholders and individual capitalists running an enterprise they own. In the 19th century capitalism an individual capitalist invested in a process of production and retained control over it. A genesis of the corporate form of business organization separated ownership and control, as was first stressed by Berle and Means (1932). Ownership has become a basis of sharing in the gains of enterprise, and control – “the power to determine the broad policies guiding a corporation” (Kotz 1978, p. 15). The separation of ownership and control poses a question: Who controls a modern business enterprise? There are three answers to this question – managers, owners, and financial institutions (Kotz 1978, p. 2-12).¹⁰⁷ For the present purposes, what matters is that a lack of control over business enterprise by the ultimate shareholders is a basis for their accepting a rate of return on investment below the profit rate, in spite of the shareholders having in principle the same property rights as an individual capitalist before incorporation. It is therefore a lack of control that forms a basis for a wealth transfer in the form of founder's profit.

Founder's profit reveals the role of control in capitalism, although in a new form peculiar to the corporate form of business organization. In the 19th century capitalism, with its unity of ownership and control, workers' lack of control over means of

¹⁰⁷ Kotz (1978) emphasized the financial control over corporations by banks – an argument particularly important for understanding the grounds on which banks share in founder's profit. One can argue, in addition to banks, managers and majority shareholders also have some control over business enterprise.

production was a foundation of a transfer of a part of value created by them in favor of capitalists (surplus value). With a separation of ownership and control, shareholders' lack of control over decision making and cash flows of a corporation has become a foundation of wealth redistribution through differences in the rates of return. In capitalism in general, control enables profit making. In a corporate form of capitalism, control enables extraction of profit of a particular type, namely, founder's profit which is shared among all the agents and institutions exercising control over corporation. A peculiar characteristic of founder's profit is that it hinges on control over enterprise at a moment of financial transaction – selling shares through an IPO. Retaining control over time is irrelevant for an extraction of this gain. Therefore, control gets divorced from its usual meaning of a medium- to long-run phenomenon, which was typical of the 19th century capitalism.

Control becomes a basis of extraction of a particular form of profit (founder's profit) and of a broader economic and political power. Control over enterprise allows for three concrete mechanisms of extraction of founder's profit – a difference in power resulting in a difference in the required rates of return, fraud, and information asymmetry. An opportunity of extraction of founder's profit, with its lump sum character preventing a repeated relationship between insiders and outsiders, makes the corporate form of business organization particularly susceptible to fraudulent practices. On the other hand, an emphasis on control as the root source of founder's profit reveals that although fighting with fraud and asymmetric information could reduce founder's profit, it would not eliminate it, as long as the difference in power of insiders and outsiders remains intact.

Founder's profit represents a redistribution of the investors' assets on the basis of the capitalist's control over enterprise. Thus, insiders "lock in the gains" at the expense of the outsiders, as argued by Lazonick (2007, p. 1006, 1010), through having a privileged access to new shares and selling them at a gain.¹⁰⁸ Rightly stressing the role of income redistribution from outsiders to insiders, Lazonick nevertheless overlooks the possibility of redistribution of stocks of monetary wealth which has a broader social significance than just a redistribution of the currently produced output. In spite of this relevant omission, this approach leads Lazonick (2007, p. 1011) to a powerful conclusion that the insiders' behavior should be theorized not as short-termism, as often alleged, but rather as "the self-interested pursuit of personal financial gain by those in positions of strategic control".

Founder's profit generated through an IPO takes a number of forms. First, it is an excess of the funds raised through an IPO over the amount of capital invested shared among a founder and venture capitalist, promoter, investment banks, and even employees having an equity stake in the form of stock options. This is the classic case considered by Hilferding. Second, often some shares are retained by stockholders after an enterprise goes public, allowing for a subsequent realization of a part of founder's profit. In light of the empirical evidence that IPOs are often underpriced, a subsequent increase in stock prices offers an additional source of founder's profit taking the form of realized capital gains.

¹⁰⁸ Blaug (1996 [1962], p. 467, emphasis in the original) argued "the great mystery of the modern theory of distribution is why anyone regards the *share* of wages and profits as an interesting problem". His point is too strong, yet it is indicative of a real issue, namely, the rising complexity of the class structure in modern capitalism. One aspect of this rising complexity is the replacement of profit of an individual active capitalist by founder's profit associated with a rise of the capital markets, which calls for a modification of the classical political economy analysis of income distribution.

Although extraction of founder's profit is taken to be the *raison d'être* of IPOs, this form of profit making requires a conversion of illiquid sunk capital into liquid marketable securities as a means towards this end. The fact that raising funds for business expansion is a special case, and in general IPOs merely convert the already existing assets into a liquid form has important implications for the nature of liquidity provided. IPOs are about market liquidity, not funding liquidity. They are facilitated by a change in form of liabilities of a firm (market liquidity), with funding for business expansion being an occasional supplement to it (funding liquidity).

Given that one of the core functions of the banking business – both commercial and investment banking – is liquidity provision through swapping promises to pay, it makes banks well suited to facilitate IPOs. The history of industrial IPOs in the US shows that although initially they were led by promoters and brokers, investment banks subsequently took over this function to resolve the limits inherent to IPOs led by promoters and brokers.¹⁰⁹ In particular, investment banks provided the foundation of trust which in turn was necessary for raising funds.¹¹⁰

The continuity between commercial and investment banking comes from the same core function of liquidity provision and from both forms of banking emerging to resolve the limits of the spontaneous liquidity provision among industrial and commercial capitalists themselves. Moreover, both types of banks initially emerged to provide market

¹⁰⁹ This process has a historical and logical parallel in bank credit emerging to resolve the limits of commercial credit.

¹¹⁰ After 1900 the distinction between commercial and investment banking in the US became blurred, as big commercial banks formed securities affiliates to handle the floating of new shares and bond issues (Kotz 1978, p. 35-39). The commercial banks' share of the business grew in the 1920s, until it was banned by Glass-Steagall.

liquidity. Commercial banks did it through substituting bills of exchange and IOUs by cash or banknotes on the asset side of capitalists' balance sheets. Investment banks did it through substituting own capital by equity shares on the liability side of capitalists' balance sheets.

Behind this continuity there is, however, discontinuity. Over time, commercial banks came to specialize in credit provision, thus, in funding liquidity. They do it by lending backed by collecting temporarily idle hoards and by creating own liabilities. In both cases, funding liquidity provided by banks involves their acquiring liabilities of the same amount. By contrast, investment banks specialize in provision of market liquidity. Thus, investment banks acquire liabilities equal to the amount of liquidity provided only for a short time period between they underwrite securities and distribute them among investors. They also acquire liabilities to hold inventories of securities to be able to make markets in them.¹¹¹

The idea that investment banking is a form of banking in some ways similar to commercial banking is recognized not only by a Marxist theory of banking, but by all approaches grounded in the bills view. For instance, it is not always appreciated that Minsky (2008 [1986], p. 256) also stressed that “the fundamental banking activity is

¹¹¹ This take on the difference between commercial banking and investment banking as the difference in the type of liquidity provided substantially differs from a common argument about investment banks not being really banks. This approach, which possibly took its roots in the legal separation between the two forms of banking initially proposed by Brandeis in 1914 and institutionalized by the Glass-Steagall Act in the US in 1933, has got its second birth in the aftermath of the recent crisis and took the form of discussions about the need to recreate this separation. According to a view developed here, such an approach is misleading, because it mistakes the legal differences that are furthermore specific to a particular country in a particular time period for the differences in the economic content of the two forms of banking. This does not deny the significance of the difference between the two institutions that arises from commercial bank deposits being insured by the government. The existence of deposit insurance makes commercial banks more likely to take on higher risks, including the ones associated with investment banking activities, than they would have in the absence of such insurance.

accepting, that is, guaranteeing that some party is creditworthy”. He was very clear that “the line between commercial banks, whose liabilities include checking deposits, other depository thrift institutions, miscellaneous managers of money (like insurance companies, pension funds, and various investment trusts), and investment bankers is more reflective of the legal environment and institutional history than of the economic function of these financial institutions” (Minsky 2008 [1986], p. 249). More recently, this idea was restated and further developed by Kregel (2010, p. 7), for whom banking business is a unity of safekeeping of deposits and liquidity creation for borrowers. On these grounds he concludes that both commercial and investment banks provide liquidity albeit do it in different ways. Hence, these two forms of banking are “functionally equivalent” (Kregel 2010, p. 9).

The change in the function of banking associated with provision of market liquidity, as opposed to funding liquidity, has important implications for the character of bank revenues and profits. When a bank provides credit and holds the corresponding asset on its balance sheet, its revenue comes from the cash flows of the borrower. An important common feature of discounting as the rudimentary form of banking and commercial banking as its more developed form is a character of profit. Namely, in both cases bank revenue arises from interest payments made by the ultimate borrower. These revenues are then used to cover the operating costs of the banking business and to pay interest on bank liabilities, with the remainder constituting bank profit. Thus, bank profit arises from income of the ultimate borrower. For example, when a borrower is commercial or industrial capitalist, bank profit comes from subdividing commercial or

industrial profit, respectively.¹¹² Incidentally, such a take on the source of bank profit does not make a Marxist theory unique. A similar understanding can also be found in both heterodox and mainstream theories that treat interest paid by banks as a form of sharing in revenue. For example, Schumpeter (1961 [1911/1934], p. 158, 201) argued that interest on demand deposits should be understood as banks' sharing their premium with depositors – a view later adopted by Bossone (2001b, p. 2266, 2003, p. 156) and Parguez (2004, p. 32). From a mainstream perspective, this view is supported, for instance, by Diamond and Rajan (2000, 2001b, 2001a). Obviously, this approach is also reflected in accounting practices and income statements.

By contrast to commercial banking, when a bank facilitates an IPO, it shares in the founder's profit. Thus, bank fees from underwriting equity shares represent a particular form of financial profit due to the gap between the two sets of prices of capital and, therefore, in the first instance come from redivision of monetary assets of the ultimate investors. It was argued above that control over enterprise forms a basis for extraction of founder's profit. Kotz (1978, p. 8, 20, 26) stresses banks' control over sources of capital forms the original basis of financial control of banks over enterprise.¹¹³ Therefore, banks' control over IPO can be seen as a basis for banks' sharing in founder's profit. A take on founder's profit as resulting from the gap between the two sets of prices of capital goods can shed some light on bank revenues for secondary equity issues and bond flotation. From a firm's perspective, acquisition of this type of liabilities, unlike

¹¹² Considering lending to industrial and commercial capitalists, Marx treats interest as a form of subdivision of surplus value, which makes interest received by banks a particular case of this subdivision.

¹¹³ Kotz (1978, p. 20, 64) stresses that the main basis of financial control of financial institutions over nonfinancial corporations has shifted over time from supply of capital in the late 19th century to holding of stock in 1946-74.

IPOs, is not driven by the motive of founder's profit extraction. It is rather used to raise additional funds (funding liquidity), or change capital structure of an enterprise by substituting one form of liabilities by another (market liquidity). Nevertheless, from a bank's perspective, floating bonds and secondary equity issues are similar to IPOs. They involve the same set of skills on the part of banks and, therefore, result in the same form of profit from the macroeconomic perspective, namely, profit from redistribution of monetary assets of the ultimate investors.

The transition from commercial to investment banking therefore involves a transformation of the character of bank revenues that is more profound than the related transformation of bank function that at its core remains liquidity provision. This "paradox of the turn" – from commercial to investment banking – explains why the social significance of this transformation cannot be fully captured by a focus only on the functions provided by banks. The character of bank revenues can be a useful lens for a more complete understanding of this transformation. A Marxist theory of banking discussed in chapter 4 is well suited to capture and explain this paradox of the turn. Its being grounded in the bills view allows it to recognize the continuity of the banking functions, while its emphasis on the sources of profit that comes from the classical political economy makes this theory of banking sensitive to the transformation of the character of bank revenues and profits.

5.7.2. Profit from Mergers and Acquisitions (M&A)

A more developed form of profit associated with a gap between the two sets of prices of capital is profit from mergers and acquisitions. Mergers and acquisitions are

buying and selling of bundles of capital assets and market positions as embodied in firms and parts of firms. Historically the first merger wave swept the US at the same time as incorporation, at the end of the 19th century.¹¹⁴ Underwriting should however be seen as a logical origin of mergers, because, first, a merger involves issuing equity shares as claims on future revenues of an enterprise, with a difference from an IPO being that this enterprise is a newly created entity consisting of at least two previously independent firms. Second, even though historically IPOs and mergers developed at around the same time, and some mergers were among non-incorporated enterprises, the majority of modern mergers are among the already incorporated businesses. Finally, from a bank's perspective, at least in the US, historically banks came to facilitate industrial mergers only after having facilitated industrial IPOs and recapitalizations (Navin, Sears 1955, p. 124). In fact it took some time for the investment banks leading railroad underwriting to underwrite securities financing the mergers, which they started doing only after having witnessed successes of individual promoters and to some extent to resolve the limitations faced by mergers led by promoters. Investment banks came to play an important role in railroad consolidation in the 1890s, and in industrial mergers – only after 1898, with J.P. Morgan's formation of Federal Steel (Kotz 1978, p. 32). For all these reasons, an analysis of incorporation and market for claims emerging with it should precede an analysis of mergers. Purchasing individual equities and bonds lays the foundation for purchasing entire lines of business, which should be reflected in the second set of prices. Minsky (see also Minsky 1993, p. 11, 1986, p. 348) was well aware of this fact, when he argued that

¹¹⁴ Incorporation started in the US with the railroad companies and predated industrial mergers.

“the starting point for bids on existing firms is the market valuation of the equity and debt liabilities”.

There are several approaches to why M&A exist, not the least because different participants of the deal – manufacturers, financiers, investors – enter it for different reasons and with various motives. The dominant view explains the existence of M&A through the benefits of the market for corporate control as an institution determining investment strategies. This view was for the first time articulated by Jensen (1986, 1989).

¹¹⁵ Alternatively, second, M&A are treated as a mechanism for controlling or alleviating competition. Third, M&A can be driven by a desire to cash in by the stakeholders of the acquired firm. This is, for instance, an explanation of the merger wave in 1895-1904 in the US by Lamoreaux (1985) who argues that stakeholders wanted liquidity as they tried to escape intensified price competition. This argument is analogous to one of the two common explanations of IPOs. Fourth, merger can allow for rationalization of production resulting in higher future profits.

Although these approaches to the existence of M&A capture some of the specific motives behind the merger deals, it is useful to shift the focus to profit extraction as a more general *raison d'être* of M&A. Mergers can allow for an increase in revenues in the future. More important however is the ability of M&A to yield profits now on the basis of an anticipated future gain and thus to lock in a gain, no matter whether the expected future gain will actually exist. Chandler (1977, p. 332) emphasizes the profit motive behind the historical origin of mergers arguing that “both manufacturers and financiers quickly learned how to profit from the actual process of legal consolidation” that swept

¹¹⁵ For a compelling critique of this approach, see Lazonick (1992, 2007).

the US in the 1890s. An important aspect of mergers is that they “often brought short-term gains” regardless of their being primarily driven by this gain or attempts to control competition (Chandler 1977, p. 339). A significance of the profit motive in the historical origin of mergers is obvious in the case of formation of the United States Steel Corporation. J.P. Morgan carried out the deal by consolidating three basic steel companies to form Federal Steel and subsequently buying out the Carnegie Steel Company and most of the remainder of the steel industry in the US. While the tangible assets of United States Steel were estimated at \$682 million, \$1321 million of stock and bonds were sold to the public, resulting in lucrative founder’s profit (Kotz 1978, p. 32-33). According to Lazonick (1992, p. 466), the motive of profit extraction also drove the conglomeration movement in the US in the 1960-70s.¹¹⁶ Similarly, if leveraged buy-outs (LBOs) in the first half of the 1980s were defensive – to protect companies from the market for corporate control – in the second half of the 1980s they became offensive and were also motivated by profit extraction (Lazonick 1992, p. 472).¹¹⁷ In both cases the market for corporate control has become a source of financial profit for managers (Lazonick 1992, p. 464, 473) and other insiders. This is in line with an argument proposed by Shleifer and Vishny (2003, p. 296) for whom mergers are “a form of arbitrage by rational managers operating in inefficient markets”. An important aspect of Shleifer and Vishny’s theory is that both the acquirer and the target managers profit by getting rid of overvalued equity. There is plenty of empirical evidence suggesting these

¹¹⁶ In this context it is not surprising the conglomeration wave of the 1960s resulted in dissolution of a third of the new conglomerates in the 1970s (Lazonick 1992, p. 461, 469).

¹¹⁷ Thus, paradoxically, the very mechanism of control of the market for corporate control later became a mechanism of reinstatement of this same market (Lazonick 1992, p. 473).

gains come at the short-term expense of shareholders of the acquiring firm and possibly also long-term expense of shareholders of the targeted firm and other stockholders.¹¹⁸

A representative critique of an approach to M&A through the lens of profit extraction can be found in Lamoreaux (1985, p. 114) who argues that "it is unlikely, therefore, that many consolidations were formed solely for the purpose of promoters' profits; real economic hardship underlay most manufacturers' willingness to sell out". But even her approach can be nested as a particular case of the profit-driven behavior, therefore, her criticism of the promoter's profit argument is unnecessarily strong.¹¹⁹ At the heart of her argument is not so much the motive to cash in as such, but rather the underlying competitive pressure and associated threat of losses. Viewed from this angle, her main thesis acquires a different meaning and becomes a particular case of a profit-driven behavior, the opposite side of which is loss aversion.

Gains from M&A have some similarity with founder's profit. In both cases a financial reorganization results in a new enterprise whose market valuation increases. In the case of M&A, a gain would exist whenever the present value of the combined enterprises exceeds the sum of their individual net present values. There is, however, an important difference between founder's profit and gains from M&A. Founder's profit arises due to shareholders' lack of control over enterprise, which translates into differences in the rates of return, with the gain often being magnified by information

¹¹⁸ See, e.g., Martynova and Renneboog (2008) for a review of empirical studies.

¹¹⁹ By the same token, Lamoreaux (1985, p. 116) underestimates the role of the capital markets in the merger wave by treating them as auxiliary, as opposed to the driving force behind the mergers. Such an approach misses the point that without the capital markets the merger wave would not have been possible in principle. Before Lamoreaux, Kotz (1978, p. 24) offered an effective counterargument to perspectives akin to hers stressing that "in an abstract economic model it is often assumed that purchases and sales take place effortlessly and directly. In actual history complex institutions may develop to facilitate market transactions".

manipulation and fraud. There is no change in the expected stream of future profit. By contrast, gains from M&A stem from an increase in the expected future profits of the two enterprises combined. This can be due to economies of scale, eliminating inefficiencies, achievement of market power as seller and/or monopsony power as buyer, economies of vertical integration, benefits of access to free cash flows of another enterprise, and other possible advantages. In spite of the difference between founder's profit and profit from M&A, an important similarity is that both gains come from an increase in the current valuation of a reorganized enterprise, which opens an opportunity for a lump-sum gain. This gain accrues to institutions and individuals having control over the enterprises under reorganization and to institutions facilitating this process (e.g., investment banks). Coming from the monetary assets of the buyers of securities of a reorganized enterprise, gain from M&A also constitutes a form of "profit upon alienation".

Similar to underwriting, mergers and acquisitions require liquidity. The targeted firm demands liquidity either in the form of cash or explicit or implicit guarantee of liquidity of the equity shares with which the targeted firm is purchased. This makes banks well suited for the business of facilitating M&A through providing access to credit for the acquiring firm and through creating the foundation of trust necessary for distributing the shares. This function performed by investment banks today also corresponds to the actual historical origin of banks' involvement in the business of M&A. For instance, Navin and Sears (1955, p. 137-138) emphasized the inherent weaknesses of the mergers led by individual promoters thus forcing banks to enter the business:

"By the very nature of their operation the merger promoters did not provide the funds necessary for an expanding industry and they could not, because of their limited capital resources, guarantee to the issuing companies the funds which a new issue of securities was intended to raise. The financial men who had been

accustomed to raising new funds on a guaranteed basis were the investment bankers, and it was to them that industry eventually turned”.

Investment banks are however not merely passive facilitators of M&A. Already in the late 19th century investment banks’ control over sources of capital has resulted in their control over nonfinancial businesses which in turn allowed investment banks to become active promoters of business consolidation (Kotz 1978, p. 30, 140). In 1898-1903 and in the 1960s investment banks played a crucial role in encouraging, advising, planning, and carrying out industrial mergers. This is not surprising in the context of theories of banking discussed in chapter 4 viewing banks as active enterprises as opposed to passive intermediaries.

On the grounds of facilitating M&A banks share in the total gain associated with M&A. Bank revenues from M&A can therefore be seen as a form of profit analogous to founder’s profit. These revenues come in the form of fees to finding merger partners, for providing advice regarding the deal and financing the merger, and for managing tender offers. Fees for M&A are an important source of revenues for investment banks. For example, in 2003 in the US investment banks facilitated M&A transactions worth \$386 billion and received fees that exceeded \$596 million (Walter, Yawson & Yeung 2008, p. 342). In 2008 investment banks received \$44 billion worldwide in fees for facilitating M&A, with \$12.5 billion captured by six largest investment banks in the US, led by Goldman Sachs at \$3.1 billion (DiNapoli 2008, p. 6). Minsky (2008 [1986], p. 220) seems to be aware of the peculiar character of gains from M&A when he argued that “the income of Wall Street operators in the mergers and acquisitions business is part of national income and output. Thus, while the meaning of the above is clear, it is not technically correct. In our type of economy, there is a peculiar output called mergers and

acquisitions... The “golden parachutes” of the 1980s yield “incomes” that are hard to relate to the standard economic view of incomes”.

5.7.3. Securitization, or Investment Banking for Banking

Securitization has attracted a lot of attention, especially over the past few years when it has often been presented as a major change in banking. The literature has been exploding with descriptions of fraud and other malpractices involved in it, in many ways bringing back the spirit of Veblen. Nevertheless, these important and valid concerns do not go to the heart of the matter.

Securitization is a process of bundling loans and selling them to special purpose vehicles (SPVs), bookkeeping entities created by banks. SPVs use these loans as collateral to issue long term bonds sold to institutional and individual investors. From a bank’s perspective, securitization is one of the many forms of funding the banking business. Unlike deposit taking and other forms of liability management, securitization results in removing loans from the bank balance sheets. For the ultimate borrower, securitization is not associated with a qualitative change in credit relations, as the borrower still has to repay the principal and the interest to the loan originator who would subsequently distribute the collected funds to the bond holders. Nevertheless, securitization is normally accompanied by lower borrowing costs and in this sense is beneficial for the borrower. For the ultimate lender, holding an asset backed security (ABS) issued by SPV is equivalent to holding any other security.

There are two forms of revenue associated with securitization, namely, first, profit from the loan sale due to a rise in the value of the loan and, second, profit from servicing

loans. A number of empirical studies have established a growing size of securitization income. For instance, Inklaar and Wang (2007, p. 33-34) showed that securitization income increased as a share of the US bank output from 1.2 per cent in 1990 to 11.4 per cent in 2004. The average annual real growth rate of securitization income also rose from 0.6 per cent in 1987-1995 to 4.1 per cent in 1995-2004. In the second time period it became the fastest growing fee-generating activity of commercial banks, followed by investment banking with its growth rate of only 1.1 per cent. The share of securitized loans went up from 10 percent to 40 percent of total loans. Ashcraft and Steindel (2008) used another methodology¹²⁰ of imputing bank output and found that securitization revenue amounted to 10.63 and 11.95 per cent of the US bank output in 2001 and 2006, respectively. Securitization output of the US bank holding companies constituted 14.65 per cent of their total output in 2006.

This rise in securitization income poses a question of what this rise signifies. From the mainstream perspective, bank profits from securitization are explained as a form of remuneration for the service provided by banks. A bank originating a loan receives profit from the loan sale due to a rise in the value of the loan. This profit is explained through the benefits of diversification. An SPV also receives profit for servicing a loan, namely, collecting payments from the ultimate borrower and passing them onto the ultimate lender.

This approach represents an accurate description of the specific services provided by a bank and an SPV. Nevertheless, it leaves two important issues unaddressed. First,

¹²⁰ Ashcraft and Steindel advocate for the usual methodology of imputing bank output that, unlike Basu, Inklaar and Wang (2008) and later studies based on their approach, treats risk taking as a service provided by banks.

under which circumstances do the services provided in the course of securitization become possible? Second, what are the macroeconomic sources of profit from securitization? For instance, do they differ from the macroeconomic sources of, say, net interest? These questions usually lie beyond the scope of the prevalent approaches to securitization, hence, answering these questions requires a different theoretical framework.

Minsky has important insights that can shed some light on the issue. He noticed that securitization is associated with a fundamental transformation of the nature of bank profit. He was very clear that “in securitization, the underlying financial instruments [such as home mortgage loans] and the cash flows they are expected to generate are the proximate basis for issuing marketable paper. Income from paper (cash flows) is substituted for the profits earned by real assets, household incomes, or tax receipts as the source of the cash flow to support the paper pledges” (Minsky 2008 [1987], p. 4). Minsky has not elaborated on this statement, and the subsequent heterodox literature on securitization has not addressed the important question of the related change in the bank profit. This is obviously not to deny that this literature has noticed other aspects of the transformation of banking associated with securitization.

The central bankers and regulators (e.g, Bernanke 2008, Knight 2008, Mishkin 2008) often identify the ‘originate-to-distribute’ view as one of the dominant factors of the crisis that started in 2007. The usual argument stresses a decline in the underwriting standards due to incentive problems of the originators whose revenues depend on the volume of the loans extended and not on their quality. If in line with the information-theoretic approach banking is understood as the business of solving information

problems, the shift to the ‘originate-to-distribute’ model undermines the core of the banking business. As a response to that, there have emerged approaches to banking aiming to show that securitized banking is still banking at core. This argument is usually based on viewing banking as a business of creating informationally-insensitive debt (Gorton 2009, p. 3-4, 14), or of credit, maturity and liquidity transformation (Pozsar et al. 2010, p. 8).¹²¹ Which of the two approaches is more adequate? Both, as each perspective has a grain of truth.

On the one hand, the mainstream theorists are right that shadow banking is still a form of banking, although for reasons other than the ones they propose.¹²² It was established above that banking should be understood as a unity of liquidity provision and an exchange of promises to pay. Whether loans are securitized or not, banks still provide liquidity for the ultimate borrowers. The shadow banking system also provides liquidity to the ultimate lenders – holders of ABS – through making markets in these securities. Therefore, shadow banks do perform the function of liquidity provision for both borrowers and lenders. Similar to the underwriting function of investment banks, in the case of securitization, the exchange of promises to pay is replaced by a promise to make markets, reinforced by the credit rating agencies that function as trust enhancers. Finally, in some cases, securitization involves the exchange of promises to pay proper. This happens when banks provide contingent guarantees to buy back an ABS if its market price falls below a certain level. In this sense the argument often found in the literature that the originate-to-distribute model undermines the essence of banking business is

¹²¹ For a critique of the “originate-to-distribute” view, see Gorton (2008, p. 67-75).

¹²² The analytical strengths of a Marxist take on banking compared to other approaches to banking, including those by Gorton (2009) and Pozsar et al. (2010), were discussed in a greater detail in chapter 4.

misleading, because it is based on a conflation of a particular form of banking with banking in general. Hence, this argument also cannot provide a strong critique of the mainstream thesis that shadow banking is still a form of banking.

On the other hand, the changes in banking noticed in the heterodox literature are important. However, in order to retain the rational kernel of the approaches focusing on these changes, it is necessary to shift the axes of the discussion. Similar to the “paradox of the turn” from commercial to investment banking, the role of the changes associated with securitization is best retained if they are analyzed through the lens of what these trends imply for the character of bank profit and implications thereof. Approached in this way, the problem with securitization is not that it undermines the functions of banking, but rather that it is associated with a fundamental shift in the character of bank profit. The banking function of liquidity provision indeed remains relatively unaltered, as rightly stressed by the mainstream, but it can co-exist with an incommensurable transformation of the nature of bank profit, which is usually overlooked in the literature. Thus, the paradox of the turn from commercial to investment banking reemerges in the context of securitization.

The theoretical framework hinging on the dual price system of capital goods suggested above can be useful for conceptualizing profit from securitization. A starting point of analysis is to locate securitization in the development of forms of banking.¹²³

Lapavitsas (2009, p. 135) suggested that securitization can be seen as “adoption of

¹²³ The importance of locating securitization in the theory of banking through relating it to discounting as the origin of banking was also emphasized by Minsky. For Minsky (2008 [1987], p. 3), securitization is “the development of techniques to “enhance credits” without accepting contingent liabilities or the investment of pure equity funds”. This statement can be interpreted as securitization undermining the fundamental banking act of accepting contingent liabilities.

investment-banking techniques” by commercial banks, which turns “lending (to earn interest) into mediating the circulation of securities (to earn fees)” (Lapavitsas 2009, p. 136). Turner (2010, p. 13) has a related insight when he argues that securitization is not a new function of the financial sector, but rather an application of the functions of the financial sector – pooling assets, transformation of risk-return characteristics of assets, maturity transformation, and market making through trading – to a new class of assets, with a finer degree of differentiation of these functions. Securitization can be thought of as a shift from commercial to investment banking considered above, but this time applied to the banking business itself. That is, like investment banks providing liquidity to industrial capitalists by floating claims on the flows of value that would be produced by these capitalists, securitization emerged as a mechanism of liquidity provision to banks themselves by floating claims on their future revenues – interest received on loans. A shift to securitization is akin to investment banking applied to commercial banking itself. On these grounds one can use insights from an analysis of founder’s profit to understand securitization revenue.

First, like founder’s profit, securitization revenue has a structured and mediated nature. When loans are bundled and sold to investors, the revenues received by a bank and an SPV come directly from loanable capital of the ultimate investors who in turn receive a claim on future revenue. If expected payments materialize, investors’ loanable capital would be restored with an increment. In that sense, bank profit from securitization would ultimately bear a relationship to the payments by the ultimate borrower. If these payments, however, do not materialize as expected, securitization profit would come from a pure redistribution of loanable capital of the ultimate investors. Like in the case of

founder's profit in general, at the moment of accrual, profit from securitization comes from the loanable capital of the ultimate investors, as rightly stressed by Itoh (1988, p. 287). And only after investors' loanable capital being restored, founder's profit can prove to have a relationship to the future flows of value, stressed by Hilferding (1910), or remain a pure Itoh-like redistribution, if the payments do not materialize. This is why Minsky (2008 [1987], p. 4) was right to stress a fundamental transformation of bank profit in the course of securitization, although he did not recognize that this is merely a particular case of profit exhibiting a structured and mediated nature.

The difference between profit from securitization and founder's profit consists in the nature of the future flows of value that are expected to restore the loanable capital of the ultimate investor. In the case of floating shares of industrial capitalists, these future flows are the surplus value produced through exploitation of labor. In the case of securitization, these flows are interest payments and the principal associated with the bank lending. That is, unlike surplus value, these flows are further removed from the process of production and linked to the credit relations. Depending on the type of the borrower, these interest payments can come from surplus value (in the case of lending to industrial capitalists), wage income (lending to workers), or tax revenue (lending to the state).¹²⁴

Second, even when payments to the ultimate investors materialize as expected, the profit from securitization, like the founder's profit, is associated with a difference in the rates of return. In that sense even a fully mediated profit retains a predatory element of profit upon alienation.

¹²⁴ The latter is usually understood in a Marxist theory as ultimately coming from surplus value.

And indeed, empirical studies of securitization in the US and Europe confirm that the coupon on ABS is normally lower than the average interest rate on the underlying assets.¹²⁵ For instance, Xu and Fung (2005, p. 405, Table 2) showed that in 1988-2001 the ratio of the weighted average coupon rate of the MBS mortgage pool to the 30-year mortgage rate has 0.97 mean and 0.96 median, implying that on average the coupon on MBS is systematically lower than the interest on underlying securities.¹²⁶ It is consistent with the time series for Ginnie Mae ARM coupon and 30 year mortgage rate in 1986-2000 (Fabozzi 2001). The same result holds for the 2000s. As shown by Colangelo and Inklaar (2010, p. 18, Figure 3), interest margin between the 4-year ABS/MBS and 1-5 year housing loans in the euro area was systematically positive from January 2003 till June 2008, and turned negative only in 2008 – after the outbreak of the crisis. Basu, Inklaar and Wang (2008, p. 18) found that the interest margin between mortgage loans and related MBS in the US was 0.8 percent in 2002-2007 and 0.69 using a 5-year moving average.

This poses a question of the social validation of the discrepancy between the two rates of return. On which grounds do investors in ABS accept a rate of return below the interest rate charged by a bank for the underlying loan?

The usual justification is associated with, first, benefits of liquidity due to the asset being tradable, second, benefits of diversification. Investors can sell an ABS, unlike

¹²⁵ This is consistent with the mainstream theories of finance, according to which mortgage note rate exceeds the MBS coupon rate by the amount of servicing fee and guarantee fee (Bhattacharya, Berliner & Fabozzi 2008, p. 5, Wang, Uryasev 2006, p. 23). Matthey and Wallace (1999, p. 9) document that “the weighted average coupons (WACs) on the underlying loans tend to run about 50 basis points above the passthrough coupon rates”.

¹²⁶ The spread between WAC and the mortgage rate is referred to in the literature as “refinancing incentive”.

a bank that used to not be able to sell a loan. The difference in the rates of return can be thought of as a form of liquidity premium, similar to liquidity premium on equities compared to the underlying real assets. By the same token, investors acquire an instrument that is not correlated with their portfolio, hence, can explore benefits of diversification. Incidentally, given that these are the benefits for investors, not for the ultimate borrowers, it is not surprising that it is the investors' loanable capital that becomes a source of bank profit from securitization.

This type of social justification is well grounded, but it only holds from the perspective of an individual investor. Viewed from the perspective of the total social capital, securitization does not generate benefits of either diversification or additional liquidity. This reveals the conditions under which services associated with securitization are rendered necessary in an economy. It happens only when there is a gap between the social nature of capital, on the one hand, and private ownership of individual claims, on the other hand. Securitization is a means for banks to exploit this gap.¹²⁷ In the act of securitization, banks give a truly social nature to a claim by giving it a marketable form and by subsequently making markets in these claims. At the same time, this act of socialization involves a private appropriation of benefits associated with it.

The structured and mediated character of profit from securitization has an important implication. As in the case of underwriting, it is the gap between the unmediated and mediated character of profit that creates particularly favorable conditions

¹²⁷ Banks do not create transferability at the moment of floatation, but rather exploit the general social function of transferability performed by the market. It was stressed by Hilferding (1910, chap. 7) that "it is the transferability and negotiability of these capital certificates, constituting the very essence of the joint-stock company, which makes it possible for the bank to 'promote', and finally gain control of, the corporation".

for fraud, misinformation, and other bank malpractices all of which has attracted a lot of attention recently. But it also implies that it is not these practices per se that are the root of the problem, they are merely symptoms of a deeper transformation of the sources of bank profit.

To recap, securitization does not involve a fundamental transformation of the core functions of banking. Nevertheless, it is accompanied by a profound transformation of the character of bank profits. In this sense the paradox of the turn from commercial to investment banking re-appears in the context of securitization. It is not surprising, given that securitization can be understood as investment banking function provided to commercial banks themselves, as opposed to non-financial corporations and the state. In spite of different concrete mechanisms of their extraction and apparently different social justification, profit from securitization and profit from underwriting have a fundamentally similar nature. Both forms of profit have a structured and mediated character, hence, both signify banks' turn to ultimate lenders, not ultimate borrowers, as an immediate source of their profit. Furthermore, even if the ultimate lenders recover their loanable capital with an increment, banks' profit still retains an element of profit upon alienation, because it hinges on the differences in the rates of return.

5.8. Conclusion

The chapter argues that the distinct conceptualization of capital assets as having two sets of prices – the market price based on capitalization of the future yield on capital assets and their reproduction costs – offers limited insights into the patterns of investment. Nevertheless, such an approach to capital assets is relevant, because the gap

between the two sets of prices constitutes space for extraction of capital gains – profit of a peculiar character. The existence and persistence of the gap itself, however, requires an explanation. Such an explanation boils down to a strategic control of an enterprise by insiders. A difference in the rates of return – the rate of return on the capital asset itself and the rate of return required by the outsiders – is a concrete mechanism facilitating transfers of monetary assets in favor of insiders. The profit due to the gap between the two sets of prices is shown to take many forms, among which are profit from initial public offerings, mergers and acquisitions, and securitization.

It is argued capital gains as a prevalent form of profit presuppose a certain level of capitalist development. At the same time, paradoxically, this most developed form of capitalist profit represents a reinstatement of profit upon alienation characteristic of the pre-capitalist modes of profit-making.

Some of the changes in banking practices can be best understood through the lens of the character of revenues they generate, because these activities involve a much greater transformation in the nature of revenues than in functions. Therefore, it is not surprising the prevalent theories cannot capture the scale of the transformation of banking due to their focus on functions. Underwriting can be taken as a logical origin of this paradox of the turn which later reappears in the case of mergers and acquisitions and securitization.

The next chapter will use the two key ideas developed in this chapter, namely, first, of a peculiar type of profit associated with the functioning of the financial markets and, second, of the multiplicity of forms taken by these revenues, to examine the transformation of the bank holding companies in the US since the mid 1980s.

CHAPTER 6

REVENUES AND PROFITS OF THE US BANK HOLDING COMPANIES:

MULTIPLICITY OF FORMS – UNIFORMITY OF SOURCES

6.1. Introduction

The goal of the present chapter is to show that behind an apparent heterogeneity of the new banking activities lies a fundamental similarity – a turn to extraction of capital gain-like revenues. These revenues are in turn associated with wealth transfers from the ultimate asset holders, including households, in favor of banks and other financial institutions.

The chapter empirically examines a composition of revenues of the U.S. bank holding companies (from now on, BHC) in 2001-2010 and shows that behind an increase in the non-interest income stands a multiplicity of non-traditional banking activities, with the single largest category being “other non-interest income”. To go beyond such inconclusiveness of a decomposition of BHC revenues based on the accounting categories, the chapter deploys a theoretical argument developed in the previous chapter, namely, the multiplicity of forms taken by profit coming from the gap between the two sets of prices of capital assets. Chapter 5 argued profit associated with a dual price system for capital assets takes many forms, among which are founder’s profit, underwriting revenues, fees from M&A, securitization revenue, and capital and trading gains. Replacing the accounting categories by the theoretical categories from chapter 5, this chapter suggests the apparent heterogeneity of the non-traditional banking activities conceals a dominant principle behind some aspects of the transformation of banking. Almost a half of the revenues associated with the non-traditional banking activities

represents capital gain-like income, therefore, extraction of this form of revenues can be seen as a driving force behind a number of changes in banking. Capital gain-like revenues are, moreover, highly concentrated in the largest BHC, with the largest seven banks receiving 80 percent, and largest twenty banks – 90 percent of the total capital gain-like revenues accruing to the BHC sector as a whole in 2009-2010.

The social significance of the transformation of banking being driven by extraction of the capital gain-like revenues lies in their macroeconomic source. These gains come from a redivision of financial assets through differences in the rates of return. It is therefore availability of monetary wealth in the society that makes these revenues possible. The redivision takes place either immediately (redivision of an ultimate asset holder's stock of wealth), or in time (when a rise in leverage of an asset holder and an immediate redivision of the borrowed funds ultimately results in a redivision of wealth of an asset holder upon debt repayment). Therefore, the transformation of banking can be seen as a shift towards redistribution of wealth from the rest of the society in favor of financial institutions.

In this context an increase in accumulated net worth and net financial wealth (financial assets less liabilities) as a multiple of GDP signifies an increase in the pool of assets to be redivided in favor of recipients of the capital-gain like revenues. Household wealth has become a particularly important source of capital gain-like revenues, for two reasons. First, household net worth and net financial wealth have risen as a multiple of their disposable personal income to a level that exceeds a similar indicator for the economy as a whole, leaving households relatively more exposed to wealth transfers. Second, this trend is reinforced by a change in the composition of the household financial

assets in favor of non-deposit financial assets that rose from 76 percent of their total financial assets in 1973-84 to 87 percent in 2000-11. This rise in size and a change in the composition of household financial assets was enabled and sustained by neoliberalism through slashing state social programs, including a turn to private pensions, and through a shift from compromise between labor and capital to a dominance of labor by capital.

These findings suggest the existing literature is right in the emphasis on the peculiar character of revenues associated with the transformation of banking. The present chapter supplements the existing studies focusing on the fictitious character of these revenues, due to either accounting standards or risk illusion, by shifting the focus to their macroeconomic sources. It is therefore argued that although accounting practices and excessive risk taking have played an important role by boosting the profit rates in the financial sector and by creating conditions favorable to asset price cycles and a crisis, there is also a structural problem with extraction of some forms of financial profit. Capital gain-like revenues can be considered fictitious in the sense that their magnitude can change suddenly for reasons unrelated to the process of production. From the perspective of the macroeconomic sources, however, far from being illusory, these profits involve a real wealth transfer, with individual forms of revenues being particular instances of a more general problem associated with a shift of banking towards extraction of capital gain-like revenues. The main contribution of the present chapter is thus in its bringing to the surface a structural problem with capital gain-like revenues, namely, their connection to implicit wealth transfers. This in turn suggests that tackling the issues of accounting and excessive risk taking, as important as it is, is not sufficient to address the problem of the financial sector profits. Capital gain-like revenues would continue to exist

and involve wealth transfers even if accounting standards and risk taking were properly regulated.

Section 2 overviews recent contributions on fictitiousness and rent-like character of bank revenues associated with the transformation of banking. Section 3 presents the results of an empirical analysis of BHC revenues – first based on the accounting categories and then based on the concepts developed in chapter 5. Section 4 discusses the social significance of capital gain-like revenues being a driving force behind the transformation of banking by connecting them to wealth transfers enabled and sustained by neoliberalism, and evaluates the suggested approach in the context of existing literature. Section 5 concludes.

6.2. Transformation of Banking: Illusory Profits, Mirages, and Rents

A concern with a rise of financial profits and their peculiar character is a relatively new phenomenon. As recently as in 2009, Friedman (2009, p. 42) argued “an important question – which no one seems interested in addressing – is what fraction of the economy’s total returns to productively invested capital is absorbed up front by the financial industry”. Although this is too strong a statement, and a question of the size and the nature of financial profit has attracted some attention before (see, e.g., Krippner (2005) for one of the earlier contributions), this issue had not been widely debated until a more recent proliferation of such banking activities and practices as securitization, mergers and acquisitions, and trading in financial assets. The total revenues and profits of the financial system and banking are often believed to be too high. This point is most clearly stated in Reid (2008) who argues between 1998 and 2008, the U.S. financial

sector has made cumulative “excess profits” amounting to \$1.2 trillion.¹²⁸ The literature usually points out a peculiar character of revenues associated with these activities and questions the social benefits of these activities thus casting doubts on the actual value added by the new banking practices.

One strand of literature focuses on illusory profits associated with failures in the accounting system. Profits are illusory, or fictitious, when accounting standards allow banks to overstate profits by recognizing revenues associated with uncertain future as current profits. According to Kerr (2011, p. 7), “much of the activity in the banking sector is aimed at nothing more than exploiting these accounting rules to register inflated fake profits and hence convert shareholders’ equity and, in extremis, debt-holders’ and taxpayers’ funds into executive bonuses”.¹²⁹ Such an overstatement of profits occurs in a number of ways. First, uncertain future cash flows can be turned into current profits by selling credit default swaps (CDS) that would however leave an insurer exposed to potential future losses. Second, an increase in value of assets, or a decrease in value of liabilities, is treated as profit due to mark-to-market accounting, mark-to-model accounting, or due to other estimation of the net present value based on implausibly optimistic forecasts. Third, inadequate provision for expected losses can boost current profits. These accounting practices are in the interests of security owners receiving capital gains and company’s executives whose pay is linked to company’s profits.

Turner (2010, p. 14) is concerned whether financial deepening has delivered value added to the economy. To address this issue, he singles out four main changes in banking

¹²⁸ The financial sector “excess profits” are defined as profits above the rate of growth of nominal GDP.

¹²⁹ Entitling his book *False Profits*, Baker (2011) also hints toward the illusory character of the financial sector profits, although he never explains in what sense these profits were false.

(2010, p. 17-21). First, a shift from the transformation of saving into investment to a prevalence of credit to households implies that the modern credit relations facilitate mainly consumption smoothing across the life-cycle and leveraged investments in the already existing assets, including acquisition of the real estate. Second, securitization has become a new form of credit intermediation. These trends have resulted in, third, an increase in the aggregate maturity transformation raising liquidity risk. Finally, fourth, there has been a rise in trading in financial assets, an analysis of which is complicated by a limited ability to distinguish the useful functions of market making and liquidity provision from proprietary trading not performing a socially beneficial function.¹³⁰

Turner attempts to connect these changes in banking with a change in the character of bank revenues by suggesting that “it is possible for financial activity to extract rents from the real economy rather than to deliver economic value” (Turner 2010, p. 6). This argument is then explicitly applied to profits from trade in financial assets (Turner 2010, p. 40). Some of these rents represent a return on a superior knowledge of the markets.¹³¹ Turner poses an important question of the nature of bank revenues associated with the new banking activities, and his distinction between activities

¹³⁰ Turner is right, a distinction between market making and proprietary trading is hard to make, though the difficulty arises not at the level of regulatory purposes, as in Turner, but rather at a conceptual level. One way to arrive at this conclusion is by acknowledging that market makers do not trade twice – once to make profit, another time to provide liquidity. Liquidity provision is always a by-product of their pursuit of profit, be it in a form of capital gains on own account or fees from trading on behalf of clients. Nevertheless, what matters is what this conclusion is used for. One way would be to try to distinguish between the two aspects of trading as best as one can for every institutional and historical setting. This approach is implicit in Turner’s concept of “an optimal level of liquidity”. A better alternative would be, however, to not try to draw a line between the beneficial market making and proprietary trading and instead shift the focus of analysis, given that such a line cannot exist at a conceptual level. Therefore, in this study the focus is shifted to the question of the macroeconomic sources of profits associated with trading in financial assets.

¹³¹ Rajiv Sethi also treats capital gains as return on information. Although he is right that capital gains cannot be treated as a return on capital, a return on information is not a good starting point of analysis of capital gains, for they accrue to both informed and uninformed traders.

delivering economic value and rents is a reinstatement of the classical political economy distinction between profit from production and profit upon alienation discussed in chapter 4. Nevertheless, Turner's answer to the question of the relationship between the new banking activities and revenues is not convincing because he does not provide a clear criterion of what exactly distinguishes revenues from the new forms of banking (rents) from the traditional interest spread, and rents from revenues delivering economic value.

Haldane, Brennan and Madouros (2010) have a more sophisticated answer to the question of the relationship between the transformation of banking and bank revenues. Their goal is to explain a rise in the financial sector share in gross value added and profits. For them, this rise is not due to a productivity miracle in the financial sector. Instead, it is a mirage reflecting a consistent under-pricing of risk in the context of increased risk-taking. Excess returns and productivity miracle "were built on an inability to measure and price risk" (Haldane, Brennan & Madouros 2010, p. 102). Therefore, they are a "risk illusion".¹³² The argument that excessive risk taking is one of the key factors behind the high profitability of the financial sector was earlier developed by Crotty (2008). Among the other factors he discusses are a growth in demand for financial products, a rise in concentration in the financial sector, and government interventions allowing the financial institutions to reproduce themselves. For Crotty (2008, p. 174), most of the high-risk strategies adopted by banks yield non-interest income, in particular, fees from securitization, from derivative creation and trading, and gains from trade on own account.

¹³² Given that banks are in the risk business, "risk illusion is no accident; it is there by design" as banks try to come up with "imaginative ways of manufacturing this commodity [risk], with a view to boosting returns" (Haldane, Brennan & Madouros 2010, p. 106).

An argument by Haldane, Brennan and Madouros (2010) builds on the studies by Wang, Basu and Fernald (2004), Wang (2003), Mink (2008), and Colangelo and Inklaar (2010) who proposed to modify the measure of the financial sector output used in national accounting by not treating risk-bearing as a productive contribution of the financial sector. Haldane, Brennan and Madouros (2010) go beyond this literature by arguing the adjusted methodology, while being a step in a right direction, overlooks risk under-pricing by the markets and, hence, it would still yield an over-estimate of the financial sector output.

Haldane, Brennan and Madouros (2010) examine the factors behind the “excess returns” in the financial sector. First, they single out three balance sheet strategies boosting risks and returns in the financial sector – increased leverage, rising trading books generating illusory profits due to mark-to-market accounting, and writing insurance for tail events. These three strategies generate excess returns due to inability to measure and price risk. An important feature of returns from the last two strategies is that they are subsequently offset by losses (Haldane, Brennan & Madouros 2010, p. 101). For example, “trading book profits were in fact largely illusory” and were subsequently offset by trading book losses, when the risk was re-priced after the onset of the crisis. Similarly, offering tail risk insurance, while profitable in boom years, also results in substantial losses once risk materializes.

The second set of explanations of the high returns to banking is based on the structural features of the financial sector. Haldane, Brennan and Madouros (2010, p. 106) suggest looking separately at individual financial activities, because “finance is anything but monolithic”. They observe what is often perceived to be low-risk and low return

activities (e.g., asset management, securities services, retail finance) actually yield high returns – between 20 and 50 percent – probably due to risk under-pricing, price inelastic demand, and reputational equilibrium. Similarly, what is perceived to be high-risk and high return activities (commercial and investment banking) yield relatively low rate of return compared to other activities in the financial sector – only 20 percent. Nevertheless, it is still high compared to returns for non-financial corporations. Among the revenues that call for explanation are trading gains. M&A generate even more puzzling returns because while being value-destroying (Bodnaruk, Massa & Simonov 2009, Kosnik, Shapiro 1997) they nevertheless yield advisory fees for investment banks amounting to 0.5-1.5 percent of the value of the deal. Similarly, underwriting fees are as high as 3-4 percent in Europe and even higher in the US. The level and persistence of these fees remain a puzzle for Haldane, Brennan and Madouros (2010, p. 105) who consider a reputational equilibrium to be one possible explanation.

For Haldane, Brennan and Madouros, the significance of their findings lies primarily in acknowledging mismeasurement of bank profits and bank contribution to the value added. This mismeasurement is particularly apparent after the onset of the crisis. For example, during the fourth quarter of 2008, equity prices of the major global banks dropped by around \$640 billion, the world GDP and world trade declined at an annualized rate of 6 and 25 percent, respectively. At the same time, the nominal gross value added by the financial sector in the UK grew at its fastest rate on record (Haldane, Brennan & Madouros 2010, p. 88). After the crisis, the financial sector value added in the US and the UK continued to rise and reached its historical high level of 9 percent in both countries. Thus, “at a time when people believed banks were contributing the least to the

economy since the 1930s, the National Accounts indicated the financial sector was contributing the most since the mid-1980s” (Haldane, Brennan & Madouros 2010, p. 88). A need for a better measure of the true value-added by the financial sector makes an improved measurement of risks associated with financial transactions a priority (Haldane, Brennan & Madouros 2010, p. 106).

A peculiar character of some forms of bank revenues that rose to prominence in the course of financialization is also stressed by Bezemer (2012, p. 36) who noticed that “the economy becomes concerned more with capital gains than with profit”. He focuses on two forms of bank profit, namely, capital gains and revenues from mergers and acquisitions. Bezemer (2012, p. 16) distinguishes between profit from production and capital gains arguing that for assets to be traded at increased prices there should be “a growth in indebtedness of the economy”, which is not the case with profit from production. Asset trade is a zero-sum game, therefore, “for someone to make a capital gain someone else must give up income or go into debt” (Bezemer 2012, p. 15). Therefore, capital gains can exist on the macroeconomic level only if there is an increase in leverage on the economy-wide level or income transfer. At the microeconomic level the distinction between capital gains and profit from production disappears as for an individual both forms of revenue provide purchasing power. On these grounds Bezemer (2012, p. 29) treats M&A as a particular form of extraction of capital gain-like revenues, namely, as “leveraged asset trades in pursuit of capital gains and fees”. A similar emphasis can also be found in Wray (2011, p. 11-12) who stresses the principal activity of investment banks has become not underwriting but trading in financial assets generating capital gains from a zero-sum game. An appreciation of the rise to prominence

of capital gain-like revenues in banking by Bezemer (2012) and Wray (2009, 2011) probably comes from Minsky (1990, p. 69-70) who introduced the concept of money manager capitalism to describe American capitalism since the 1970s. Accumulation of monetary wealth in the previous decades and privatization of pensions resulted in accumulation of funds in the hands of fund managers, or institutional investors. Minsky emphasized the rise to prominence of the capital gains in the money manager stage of capitalism, mainly due to money managers maximizing total return on their investment that comprises dividends and capital gains.

To recap, recent literature recognizes the transformation of banking is associated with revenues and profits of a peculiar character, distinct from profits in the sphere of production. These revenues are often treated as illusory, fictitious, or fake – either in an accounting sense or due to a risk illusion. According to the first approach, bank profits are overstated due to accounting standards treating uncertain future gains as current profits. This happens in the case of mark-to-market and mark-to-model accounting, and inadequate provision for losses (Kerr 2011). A second approach attributes “excessive returns” in finance to risk under-pricing in the context of increased risk taking (Crotty 2008, Haldane, Brennan & Madouros 2010). If profits were properly adjusted for risk, returns would have been significantly lower. Both approaches focusing on fictitiousness of bank profits emphasize mismeasurement of value added of the financial sector and, therefore, of GDP as a whole. Furthermore, the illusory character of financial profits implies these “paper profits” would be subsequently offset by losses of the recipients of these profits, as it happened during the crisis that started in 2007.

An alternative though less developed explanation views bank revenues from non-traditional activities as rents extracted from the rest of the economy suggesting an income transfer without value creation (Turner 2010). This line of reasoning can be seen as a reinstatement of the classical political economy distinction between profit from production and profit upon alienation, discussed in chapter 4. The strength of this approach, compared to the literature on fictitiousness of profits, lies in focusing on the macroeconomic source of financial profits, as opposed to accounting standards and mismeasurement of risk.¹³³ This approach is however usually explicitly applied only to capital gains and trading gains (Turner 2010, Wray 2009, Bezemer 2012, Wray 2011), and when it comes to other financial activities it remains unclear whether they are a form of rent extraction or value creation.

The subsequent section empirically examines a rise in the financial sector profits and the transformation of revenues of the U.S. bank holding companies. After an analysis based on accounting categories, theoretical categories developed in chapter 5 are deployed to suggest a possible interpretation of the structure of bank revenues and their connection to redistribution of financial assets, or wealth transfers. Grounded in the classical political economy distinction between profit from production and profit upon alienation, discussed in chapter 4, this interpretation can be seen as a critical development

¹³³ Profits overstated due to accounting standards and excessive risk taking can be considered fictitious, or illusory, in the sense that they are not immediately linked to the process of value creation and can therefore change suddenly and for reasons unrelated to the process of production – due to legal changes (changes in the accounting practices) or changes in the market valuations of the assets and pricing of risk. Capital gain-like revenues is one form of profits fictitious in this particular sense. Interest has similar characteristics. By shifting the focus from fictitiousness of profits to wealth transfers, the chapter alters the social implications of extraction of this form of revenue. The social implications of fictitiousness of profits are related to the question of financial stability and redistributing outcomes arising in a crisis. By contrast, capital gain-like profits as understood in the present study bring to the surface wealth transfers across the society above and beyond the current income, taking place regardless of a crisis.

of the second approach to bank profit discussed in the present section – the one focusing on its macroeconomic sources as opposed to unreality of these profits – applied to a range of financial revenues.

6.3. Revenues and Profits of the U.S. Bank Holding Companies: An Empirical Analysis

Profits accruing to the U.S. financial sector have risen from 13 percent of total domestic profits in the 1950-60s to 32 percent in 2000-11, on average (Figure 2).¹³⁴ Commercial banks profits as a share of the total domestic profits doubled during the same time period – from 6 to 12 percent.¹³⁵ Both figures reached their historical record level in the early 2000s, when the financial sector profits rose to 40 percent and commercial bank profits – to 19 percent of the total domestic profits.

An alternative way to illustrate a dramatic rise in financial profits is with an aid of an index showing the size of financial profit in comparison to a base year (1970 in the present case, chosen to reflect the beginning of financialization).¹³⁶ Figure 3 shows that in 1970-2011 GDP has been growing steadily and reached 15 multiples of GDP in 1970. The nonfinancial sector profit growth has been fluctuating around the trend of GDP. By contrast, between 1970 and 2006, the financial sector profit increased by 28 times. It

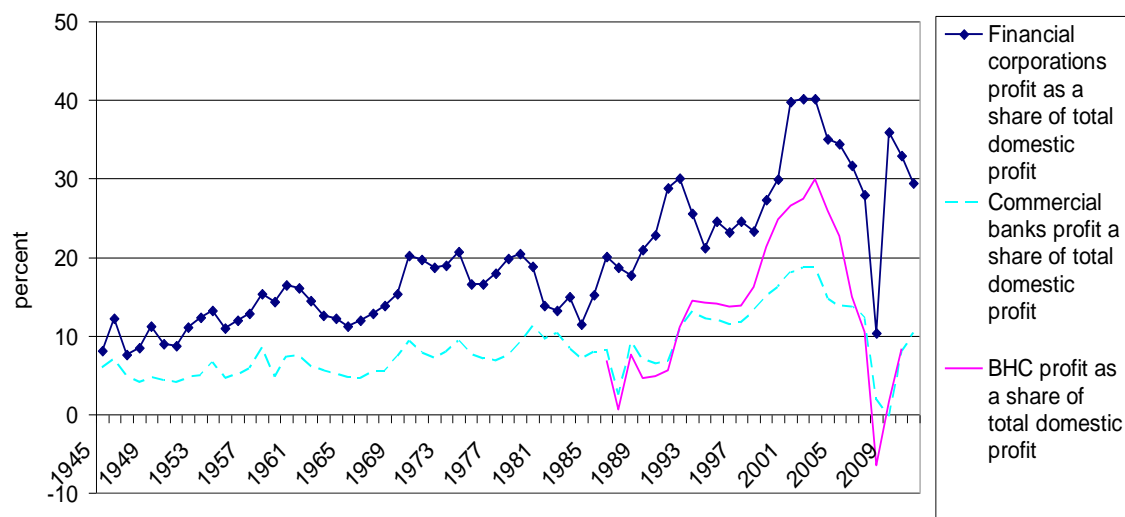
¹³⁴ Financial and non-financial sector profits are corporate profits with inventory valuation (IVA) and capital consumption adjustments (CCAdj).

¹³⁵ Commercial banks pre-tax profits are calculated as a sum of pre-tax net operating income and net securities gains (losses). Pre-tax net operating income is a sum of net interest income and total noninterest income less total noninterest expense less the provision for loan and lease losses. Since 1959, net securities gains (losses) represent the net value of profits on securities sold or redeemed less losses on securities sold. In 1941-1958, net securities gains (losses) are the net value of recoveries on securities less losses and charge-offs on securities.

¹³⁶ This form of illustrating a rise in the financial profit comes from Jim Reid (2008).

dropped in 2008, but very quickly recovered, and already in 2010-11 exceeded the pre-crisis level by having reached 31 and 29 multiples of the 1970 level, respectively.¹³⁷

Figure 2: Pre-tax financial sector profit as a share of total pre-tax domestic profit (US, 1945-2011, different sectors).



Source: calculations by author, NIPA, Table 6.16; FDIC, Historical Statistics on Banking, Table CB04; FR Y-9C Reports. Domestic profit is with IVA and CCAdj. BHC and CB profit is after provision for loan losses, with realized capital gains included.

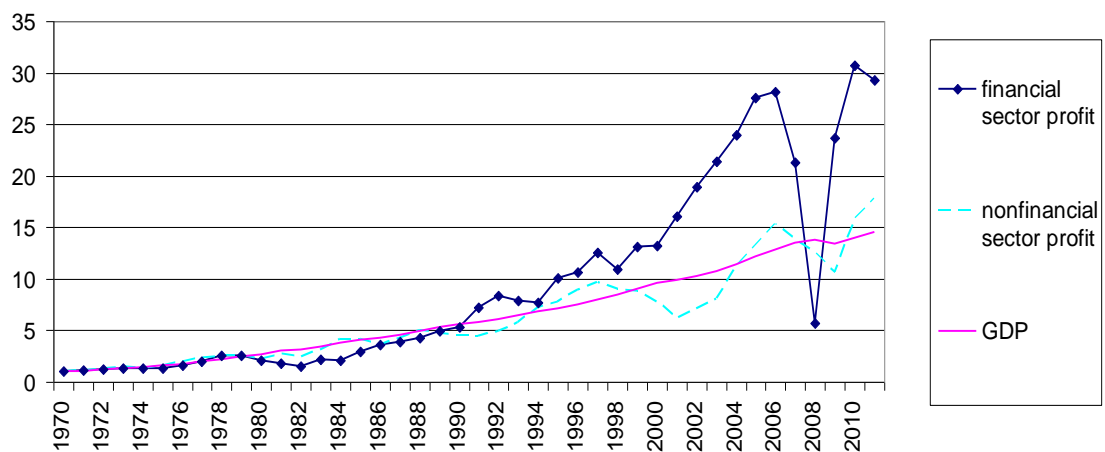
The trend in the financial sector profit as a share of the total domestic profit is closely followed by a trend in the BHC profit as a share of total (Figure 2).¹³⁸ According to the Bank Holding Company Act of 1956¹³⁹, a bank holding company is a company which has control over any bank. In December 2010, 80 percent of banks in the U.S.

¹³⁷ The quick restoration of profits would have obviously not been possible to the same extent without massive government interventions to bail out banks.

¹³⁸ In the FR Y-9C reports, BHC pre-tax profit is reported as income (loss) before income taxes, extraordinary items, and other adjustments. It is a sum of net interest income, total noninterest income, realized gains (losses) on held-to-maturity securities and realized gains (losses) on available-for-sale securities, less total noninterest expense, less provision for loan and lease losses, and less provision for allocated transfer risk.

¹³⁹ The Act is available online at <http://www.fdic.gov/regulations/laws/rules/6000-100.html>

Figure 3: Pre-tax financial sector profit, nonfinancial sector profit, and GDP (US, 1970-2011). Index: 1970 = 1.



Source: calculations by author, NIPA, Tables 1.1.5, 6.16. Domestic profit is with IVA and CCAdj.

were owned by BHC.¹⁴⁰ BHC profits rose from 7 percent in 1986 to 30 percent of the total domestic profits in 2003. Profits declined in the following years and were followed by losses in 2008, before recovering in 2009-10. Not only do BHC profits trace the general trend of financial profits in the US, but they also amount to a substantive share of the total financial profits. BHC pre-tax profits rose as a share of the total financial sector profit from the average 26 percent in 1986-1992 to 67 percent in 1993-2005, with a maximum of 83 percent in 2000.¹⁴¹ Given the relative size of the BHC profits in the profit of the financial sector as a whole, and given the similarities in the two trends, the rest of the section focuses on an empirical analysis of BHC revenues and profits. A

¹⁴⁰ Data are from <http://www.fedpartnership.gov/bank-life-cycle/charts/bank-ownership-by-bhcs.pdf>.

¹⁴¹ The relative size of BHC in the financial sector as a whole changed during the crisis, when BHC sector suffered losses in 2008, unlike the financial sector as a whole. In 2009 and 2010, BHC profits rose but only to 4 and 25 percent of the total financial profits, respectively. BHC are therefore representative of the financial sector as a whole only in 1993-2005.

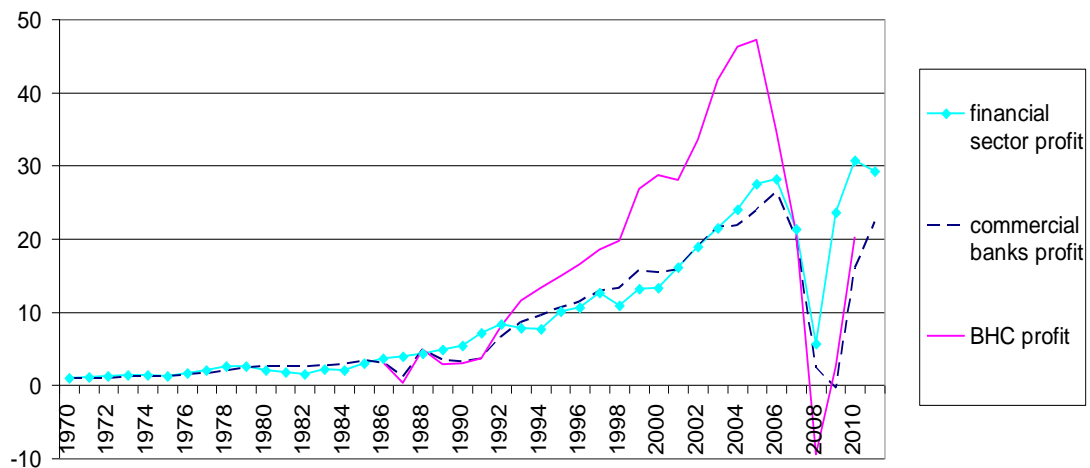
decomposition of BHC revenues and profits would shed the light not only on the ongoing transformation of banking, but also on the dynamics of the financial profit in general.

In addition to BHC being a good lens for studying the transformation of profits of the financial sector as a whole, BHC profits are a worthwhile object of analysis on their own account as they have been growing at a pace exceeding that of profits of the financial system as a whole. If by 2005 the financial sector profits rose to 28 multiples of their 1970 level, and commercial bank profits closely followed this trend, the BHC profits peaked at 47 multiples of their 1970 level (Figure 4).¹⁴² The significantly higher rate of growth of BHC profits is partly due to a faster rise of profits of the existing BHC and their subsidiaries, and partly – due to creation of new BHC and acquisition of new banks and non-bank subsidiaries by the already existing BHC. The percentage of banks owned by BHC in the U.S. more than doubled in the 1980s. It rose from 34.3 percent in 1980 to 72 percent in 1990.¹⁴³ Although since then it increased at a much slower pace, this process still contributed to an increase in BHC profits. By 2000 the percentage of banks owned by BHC reached 80.1 and in the 2000s it fluctuated around 83 percent. Regardless of the specific reasons behind a sharp increase in BHC profits, this increase indicates a rising significance of this type of financial institutions in the US, which in turn makes an examination of the structure of BHC revenues even more relevant.

¹⁴² The data for BHC are available only since 1986. To make the BHC profit index comparable with the other indices, it is assumed that in 1970-1985 the BHC profits rose at the same pace as commercial bank profits. Thus, for the BHC profit index 1986 is taken as a base year, with an index number set at 3.19 – the value of the profit index for the commercial banks that year.

¹⁴³ Data come from <http://www.fedpartnership.gov/bank-life-cycle/charts/bank-ownership-by-bhcs.pdf>.

Figure 4: Pre-tax profits of the financial sector, commercial banks, and BHC (US, 1970-2011). Index: 1970 = 1.



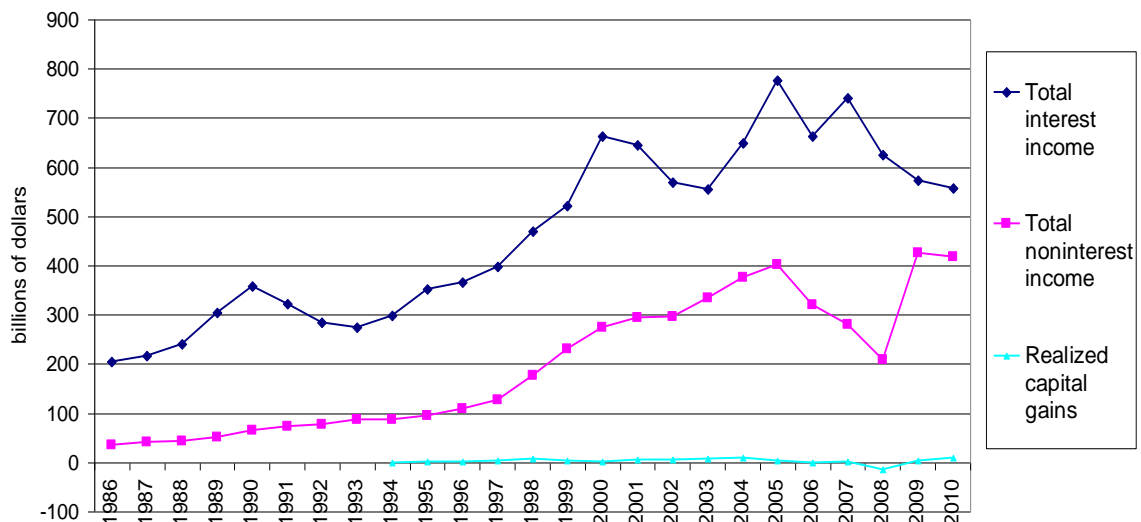
Source: calculations by author, NIPA, Table 6.16; FDIC, Historical Statistics on Banking, Table CB04; FR Y-9C Reports. Financial sector profit is with IVA and CCAAdj. BHC and CB profit is after provision for loan losses, realized capital gains included.

The data for the US bank holding companies come from the FR Y-9C Reports filled out by bank holding companies, collected by the Federal Reserve Bank of Chicago and available through Wharton Research Data Services (WRDS), The Bank Regulatory Database. The database includes income statements and balance sheets for bank holding companies for 1986-2010. For the purposes of this study, I aggregate these data to arrive at an aggregated income statement and balance sheet for the bank holding companies sector as a whole.

The usual argument about the evolution of sources of bank revenues stresses a long-term increase in the non-interest income, in absolute terms and as a share of total revenues. Figures 5 and 6 confirm it. Total noninterest income of the U.S. BHC increased from \$36 billion in 1986 to \$401 billion in 2005. After a decline to \$209 billion by 2008, noninterest income doubled and reached its historical record of \$420 billion in 2009-2010. Banking activities generating non-interest income have been growing at a faster

pace than the traditional banking business. As a result, between 1986 and 2003, noninterest income has increased from 15 to 38 percent of total revenue, and after a decline in the pre-crisis years reached unprecedented 43 percent in 2009-2010. In 1986-2010, noninterest income as a share of net operating revenue has increased from 30 to 50 percent, as well. It means nowadays almost a half of total BHC revenues and a half of revenues net of interest expenses are generated through non-interest income. Finally, after rising gradually as a share of GDP in 1986-1996, noninterest income rose sharply after 1997. As a result, between 1997 and 2004 noninterest income as a share of GDP more than doubled – from 1.5 to 3.2 percent. The composition and the nature of non-interest income, therefore, call for analysis.¹⁴⁴

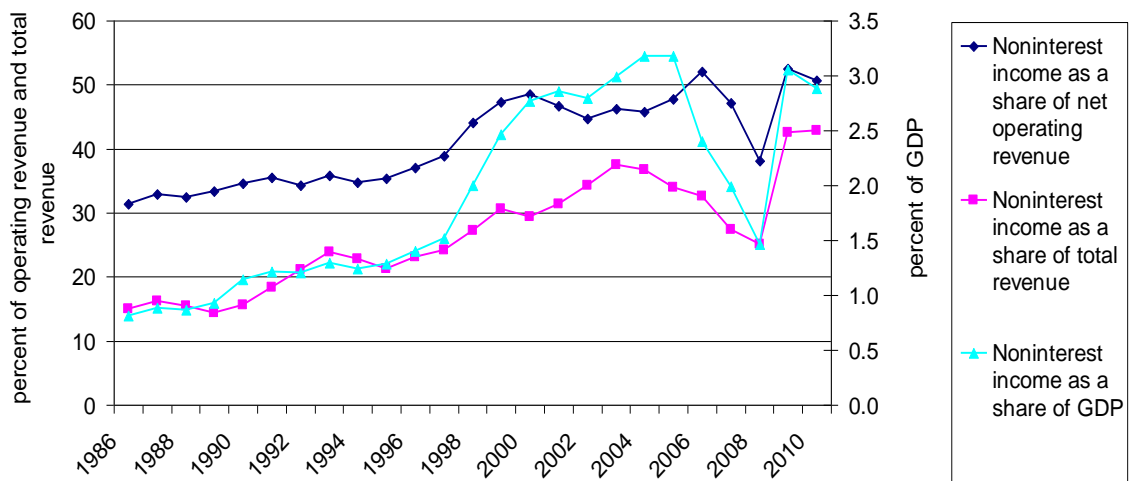
Figure 5: Total interest income, non-interest income, and realized capital gains (US BHC, 1986-2010).



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

¹⁴⁴ Another important trend in bank revenue extraction is a change in composition of the interest income. This question lies beyond the scope of the present analysis. Suffice it to say that in 2008-2010 between 18 and 20 percent of total interest income of the US BHC was coming from residential mortgage lending.

Figure 6: Noninterest income as a share of net operating revenue, total revenue, and GDP (US BHC, 1986-2010).



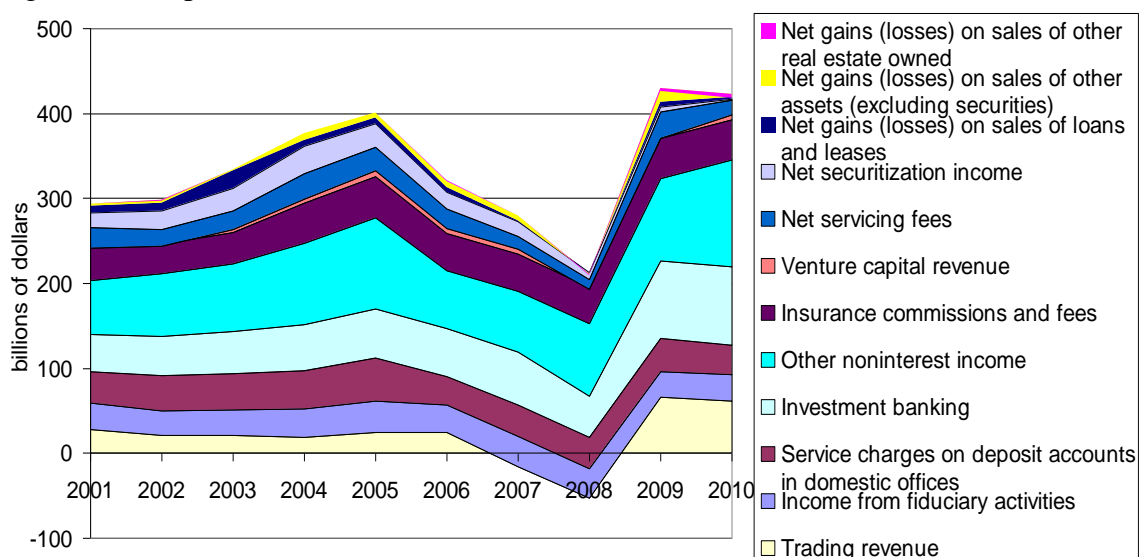
Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS; NIPA, Table 1.1.5.

A closer look at the composition of the non-interest income reveals the multiplicity of activities and bank services hidden behind it (Figure 7). Noninterest income includes trading revenue, income from fiduciary activities, service charges on deposit accounts, fees from investment banking, insurance commissions and fees, venture capital revenue, net servicing fees, net securitization income, and net gains on sales of various assets.¹⁴⁵ Thus, a decomposition of the sources of income based on the accounting categories remains inconclusive, especially given that the single largest component of the total noninterest income is “other noninterest income”.¹⁴⁶

¹⁴⁵ Prior to 2001, non-interest income was sub-divided into income from fiduciary activities, service charges on deposit accounts, trading revenue, and other non-interest income. The latter was rising and since at least 1996 accounted for more than a half of total non-interest income. Therefore, a meaningful decomposition of non-interest income useful for our purposes is only possible since 2001, when methodology was changed to further decompose other non-interest income. For this reason, the remainder of the empirical analysis in this chapter will focus on 2001-2010.

¹⁴⁶ Other noninterest income includes income and fees from the printing and sale of checks, earnings on/increase in value of cash surrender value of life insurance, income and fees from automated teller machines (ATMs), rent and other income from other real estate owned, safe deposit box rent, net change in

Figure 7: Composition of non-interest income of the US BHC, 2001-2010.¹⁴⁷



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Only two facts stand out from the decomposition grounded in the categories given in the income statements. First, the BHC sector as a whole incurred trading losses in 2007-2008 and, second, other than that, in 2001-2010 the relative shares of different types of non-interest income have not changed much. These two observations do not change the overall inconclusiveness of this decomposition of the non-interest income. The myriad of activities generating non-interest income, none of which taken individually is large enough to account for a substantial part of BHC revenues, creates an appearance of a lack of a dominant principle behind the non-traditional banking business.

the fair values of financial instruments accounted for under a fair value option, bank card and credit card interchange fees, and gains on bargain purchases.

¹⁴⁷ Net gains (losses) on sales of other real estate owned are visible on the graph only in 2009-2010, and even then they are hardly recognizable. In other years they do not appear on the graph at all due to their negligible size. In 2001-2008, net gains (losses) on sales of other real estate owned averaged to -\$0.32 billion, with a maximum loss in the period being \$1.48 billion in 2008 and a maximum gain of \$0.27 billion in 2006.

A conclusion about the heterogeneity of the new banking activities yielding non-interest income is true, however, only insofar as the concrete services and the concrete mechanisms of revenue extraction associated with these services are concerned. This apparent heterogeneity conceals an important similarity among some of these new activities, namely, their being driven by extraction of capital gain-like revenues. An appreciation of this fact allows for discerning a common dominant principle behind some of the new banking activities. This in turn allows for a shift in focus from treating non-interest income as something merely different from interest revenue to the emphasis on the social significance of the rise of the non-interest income, which would have been overlooked otherwise.

An argument about the multiplicity of forms of profit associated with the gap between the two sets of prices developed in chapter 5 offers a possible theoretical framework for understanding the significance of the non-interest income. Some of the forms of non-interest income provided in the BHC income statements can be grouped using the theoretical categories from the previous chapter. Founder's profit would then correspond to fees from investment banking activities, including IPOs, underwriting, and M&A.¹⁴⁸ Total securitization income would comprise net securitization income, net servicing fees, and net gains (losses) on sales of loans and leases.¹⁴⁹ Realized capital

¹⁴⁸ Fees from investment banking activities are defined as investment banking, advisory, brokerage, and underwriting fees and commissions for 2001-2006 and a sum of, first, fees and commissions from securities brokerage, second, investment banking, advisory, and underwriting fees and commissions, and, third, fees and commissions from annuity sales for 2007-2010.

¹⁴⁹ Net securitization income includes net gains (losses) on assets sold in securitization transactions, i.e., net of transaction costs. It includes fees (other than servicing fees) earned from the bank's securitization transactions and unrealized losses (and recoveries of unrealized losses) on loans and leases held for sale in securitization transactions. Net servicing fees include income from servicing real estate mortgages, credit cards, and other financial assets held by others. Net gains (losses) on sales of loans and leases include

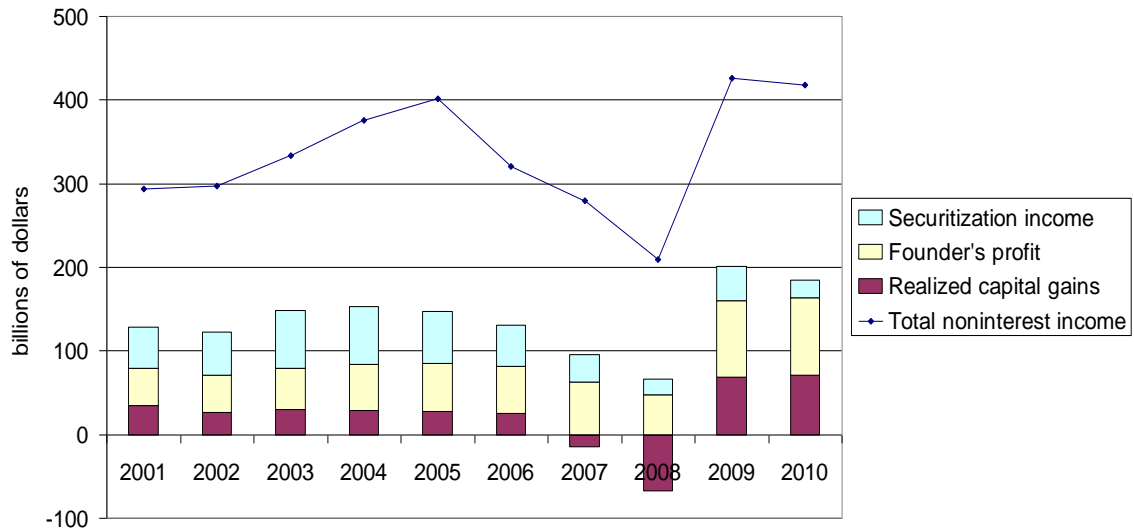
gains would include trading revenue and two types of income not being a part of non-interest income, namely, realized gains (losses) on held-to-maturity securities and realized gains (losses) on available-for-sale securities.¹⁵⁰ A sum of these three forms of capital gain-like revenues – founder’s profit, total securitization income, and realized capital gains – yields total capital gain-like revenues of BHC. In what follows below, the categories of founder’s profit, total securitization income, realized capital gains, and total capital gain-like revenues will be used in the meanings just discussed. Figure 8 shows a decomposition of BHC non-interest income based on these theoretical categories.

This decomposition brings to the surface several facts. First, in the run-up to the crisis that started in 2007, realized capital gains were the smallest revenue associated with the dual price system. For the BHC sector as a whole, trading gains were not as significant as it might appear based on how much attention they have attracted. This is probably due to the fact that some of these gains represent redistribution within the BHC sector itself, so that gains of some BHC were losses of others. In 2009-2010 trading gains more than doubled compared to their pre-crisis level of \$29 billion per year, on average. Amounting to \$70 per annum the last two years, realized capital gains became the second largest source of revenue, after founder’s profit.

unrealized losses (and subsequent recoveries of such net unrealized losses) on loans and leases held for sale.

¹⁵⁰ Trading revenue includes the net gain or loss from trading cash instruments and off-balance sheet derivative contracts (including commodity contracts) that has been recognized during the calendar year-to-date. Realized gains (losses) on held-to-maturity and available-for-sale securities are the difference between the sales price (excluding interest at the coupon rate accrued since the last interest payment date, if any) and its amortized cost.

Figure 8: Capital gain-like revenues (US BHC, 2001-2010).



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Incidentally, although “paper profits” associated with fair value accounting have been an important factor in determining behaviour of individual banks and generating leverage cycles (Adrian, Shin 2010), from the macroeconomic perspective these revenues have not been an important source of the total bank revenues in the years for which the data are available. In 2007-2010, the only years when the data are available, net change in the fair values of financial instruments accounted for under a fair value option was close to the value of realized capital gains, with an exception of 2008 (Table 1). Nevertheless, in 2009-2010 this form of income was substantially smaller than trading revenue. Revenues associated with mark-to-market accounting amounted to 0.33-3.95 percent of total non-interest income in 2007-2010.

Second, in 2001-2006 total securitization income was a far more important source of BHC revenues than realized capital gains. Averaging to \$58 billion during those years, this form of revenue was the single largest form of BHC capital gain-like revenues. In particular, it was twice as high as realized capital gains and trading gains. Although

securitization income has declined since the onset of the crisis, it still generates for BHC \$20-40 billion per year.

Table 1: Net change in the fair values of financial instruments accounted for under a fair value option in comparison to other capital gain-like revenues (billion of \$) and as a share of non-interest income (in percent).

| | 2007 | 2008 | 2009 | 2010 |
|--|--------|--------|-------|-------|
| Net change in the fair values of financial instruments accounted for under a fair value option | 0.92 | 8.26 | 2.49 | 10.65 |
| Realized capital gains | 1.50 | -14.40 | 2.89 | 10.12 |
| Trading revenue | -15.70 | -53.14 | 66.18 | 61.11 |
| Net change in the fair values of financial instruments accounted for under a fair value option as a share of non-interest income | 0.33 | 3.95 | 0.58 | 2.54 |

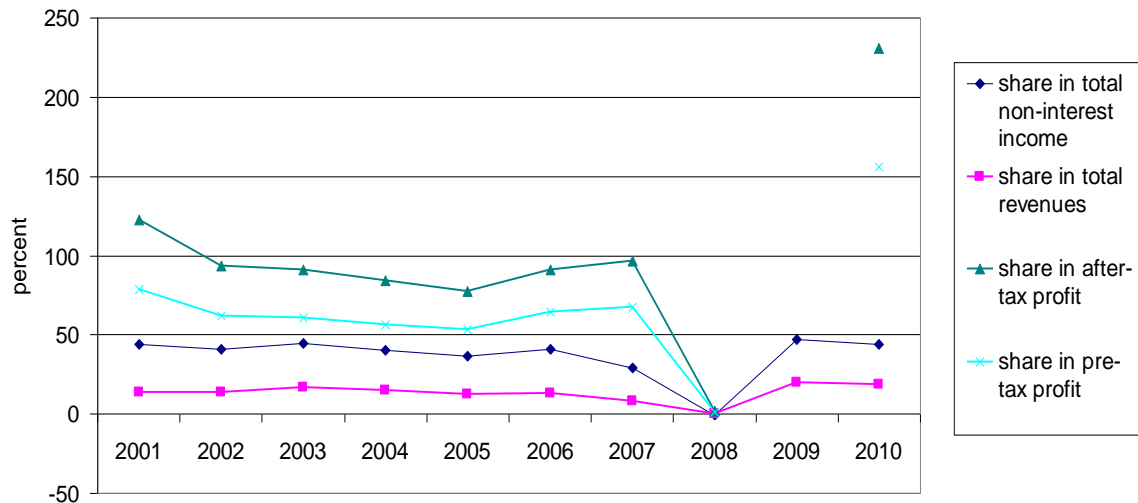
Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Finally, third, prior to 2009, founder's profit accounted for \$52 billion per year, on average, with the crisis years not being an exception. In 2009-2010 it rose to \$91 per annum and thus became the single largest form of BHC capital gain-like revenues, having taken the place of securitization income.

The significance of the capital gain-like revenues becomes apparent once it is recognized that these revenues, taken in the multiplicity of their forms, were responsible for 13 percent of BHC total revenues and for 40 percent of BHC total non-interest income in 2001-2007, on average (Figure 9).¹⁵¹ It suggests that almost a half of what is normally known as non-traditional banking activities – activities generating non-interest income – are driven by extraction of capital gain-like revenues.

¹⁵¹ Insurance commissions and fees, service charges on deposit accounts, and income from fiduciary activities are the other largest components of the non-interest income.

Figure 9: Total capital gain-like revenues, as a share of income and profit (US BHC, 2001-2010).



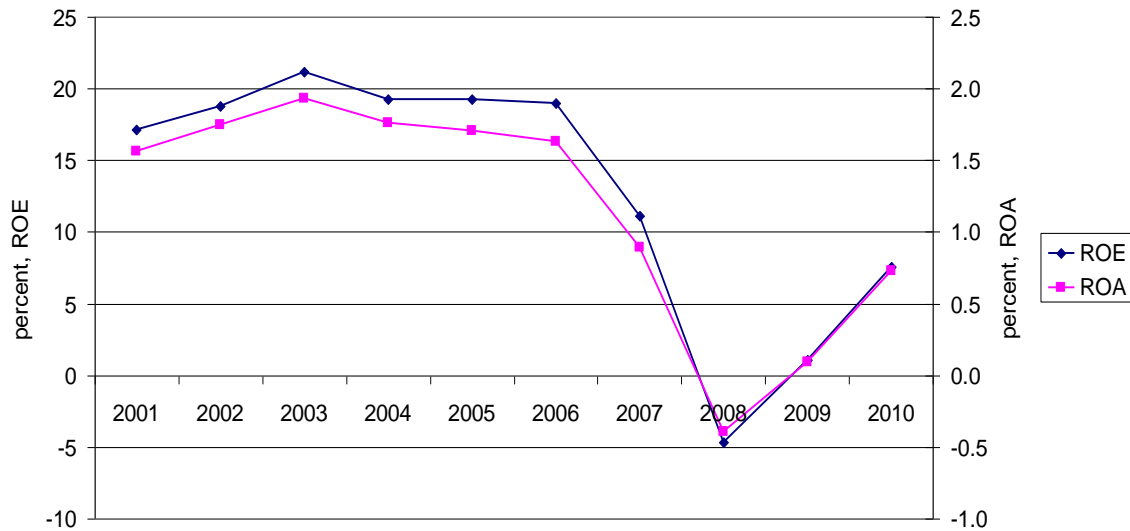
Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Furthermore, since the crisis the share of the capital gain-like revenues in total BHC revenues has increased. After a net loss from these activities combined in 2008 – a loss driven by losses from trading and capital losses – capital gain-like revenues quickly recovered and reached a historical record level of \$202 and \$185 billion in 2009 and 2010, respectively.¹⁵² As a result, total capital gain-like revenues rose to 46 percent of total non-interest income and 19 percent of total revenues in 2009-2010, on average – the relative shares also exceeding those prior to the crisis. This quick recovery of the capital gain-like revenues and their rise above the pre-crisis levels stands in sharp contrast to the fact that profitability of the US BHC still has not reached its pre-crisis level. As shown on Figure 10, return on equity (ROE), which averaged to 19 percent in 2001-2006, dropped to -4.7 percent in 2008 and rose merely to 7.6 percent by 2010. Return on assets (ROA)

¹⁵² As in the case of a quick restoration of bank profits after the crisis in general (see footnote 10), a rapid recovery of the capital gain-like revenues to a level exceeding their level before the crisis might not have been possible to the same extent if it were not for government interventions to bail out banks and provide them with other forms of assistance.

followed the same trend by falling from 1.7 percent in 2001-2006, on average, to -0.4 percent in 2008 and recovering only less than a half of its pre-crisis level by reaching 0.7 percent in 2010.

Figure 10: Return on equity and return on assets, US BHC, 2001-2010.



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Most of these findings also hold for the largest BHC, with some of them being even more pronounced for the biggest BHC than for the sector as a whole. An interest in the sources of revenue and other trends for the largest BHC usually stems from a high concentration of the banking sector. The largest 7 banks received on average 45 percent of total revenues of the BHC sector in 2001-2005 and 58 percent in 2006-2010. For the largest 20 banks the same figures increased from 67 to 78 percent in the same time period. For the largest 3, 5, 7, and 20 BHC¹⁵³, the relative size of the capital gain-like

¹⁵³ The largest BHC are defined by the size of their total revenue, i.e. a sum of their interest and non-interest income. A list of the largest BHC varies by year. In 2010, the top 20 BHC – from the largest to the smallest – were: Bank of America Corporation, JPMorgan Chase & Co., Citigroup, Wells Fargo & Company, Metlife, Goldman Sachs Group, Morgan Stanley, American Express Company, U.S. Bancorp, HSBC North America Holdings, Capital One Financial Corporation, PNC Financial Services Group, Ally Financial, Bank of New York Mellon Corporation, Taunus Corporation, State Street Corporation, BB&T Corporation, Suntrust Banks, Discover Financial Services, Regions Financial Corporation.

revenues has followed roughly the same trend as for all BHC, with a drop in 2007-2008 and a return to the pre-crisis levels or even exceeding those in 2009-2010 (Figures 11-12).

Figure 11: Capital gain-like revenues as a share of total non-interest income (US BHC, 2001-2010).



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

The share of capital gain-like revenues in total revenue is, however, increasing in the bank size. For instance, in 2001-2006, capital gain-like revenues amounted to 41.25 percent of total non-interest income of all BHC, 44.06 percent for the largest 20 banks, and 56.83 percent for the largest 3 banks (Table 2). Thus, the share of capital gain-like income in the total non-interest income of the largest 20 banks was 7 percent higher than for all banks, on average (Table 3). For the largest 3 banks, it was 38 percent higher. Similarly, for the same years capital gain-like income amounted to 15.99 and 19.20 percent of the total revenue of the largest 20 and 3 banks, respectively. For all BHC, this figure equalled to 14.11 percent, making the relative size of the capital gain-like revenues for the largest 20 and 3 banks exceeded that for all banks by 13 and 36 percent,

Figure 12: Capital gain-like revenues as a share of total revenues (US BHC, 2001-2010).



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Table 2: Capital gain-like revenues as a percentage of total non-interest income and as a percentage of total revenue, all BHC and the largest BHC, simple averages for 2001-06 and 2009-10.

| | | All BHC | Top 3 | Top 5 | Top 7 | Top 20 |
|---|---------|---------|-------|-------|-------|--------|
| Capital gain-like revenues as a percentage of total non-interest income | 2001-06 | 41.25 | 56.83 | 53.36 | 46.23 | 44.06 |
| | 2009-10 | 45.72 | 59.19 | 53.95 | 57.49 | 47.85 |
| Capital gain-like revenues as a percentage of total revenue | 2001-06 | 14.11 | 19.20 | 18.42 | 17.08 | 15.99 |
| | 2009-10 | 19.41 | 22.34 | 22.66 | 26.50 | 22.00 |

Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Table 3: Capital gain-like revenues as a share of total non-interest income and as a share of total revenue: ratio of the largest banks to all BHC.

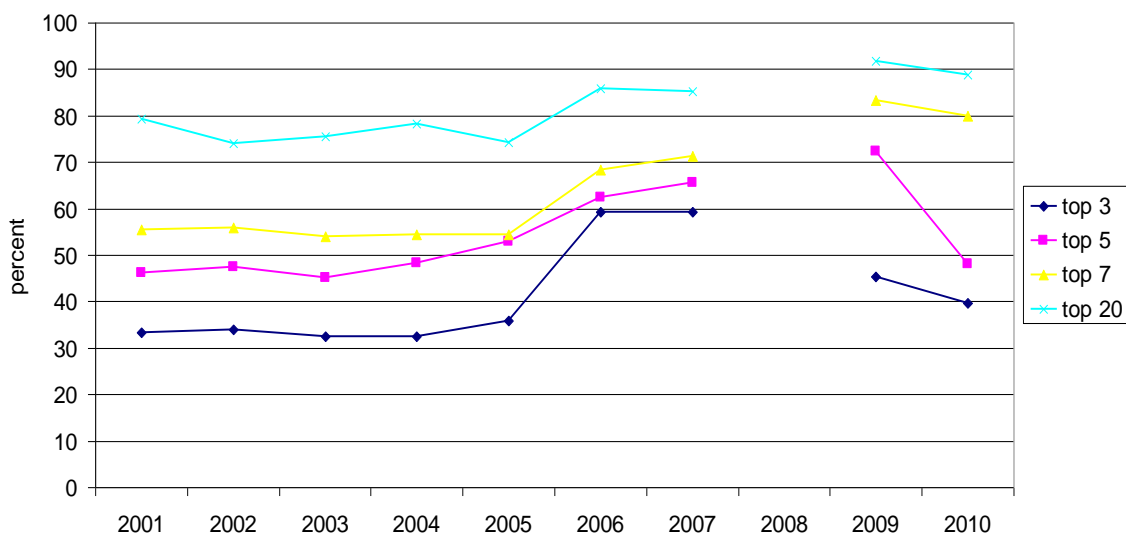
| | | Top 3 | Top 5 | Top 7 | Top 20 |
|--|---------|-------|-------|-------|--------|
| Capital gain-like revenues as a share of total non-interest income | 2001-06 | 1.38 | 1.29 | 1.12 | 1.07 |
| | 2009-10 | 1.29 | 1.18 | 1.26 | 1.05 |
| Capital gain-like revenues as a share of total revenue | 2001-06 | 1.36 | 1.30 | 1.21 | 1.13 |
| | 2009-10 | 1.15 | 1.17 | 1.37 | 1.13 |

Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

respectively. Thus, there is a tendency for the larger banks to rely more heavily on the capital gain-like revenues.

A greater reliance on capital gain-like revenues by the largest BHC results in a high concentration of these revenues in the largest institutions. Figure 13 shows the largest 7 banks received about a half of the total capital gain-like revenues accruing to the BHC sector as a whole in 2001-2005. In 2009-2010, this share increased to 80 percent. In the same time period, the relative share of the capital gain-like revenues of the largest 20 banks rose from 75 to 90 percent.

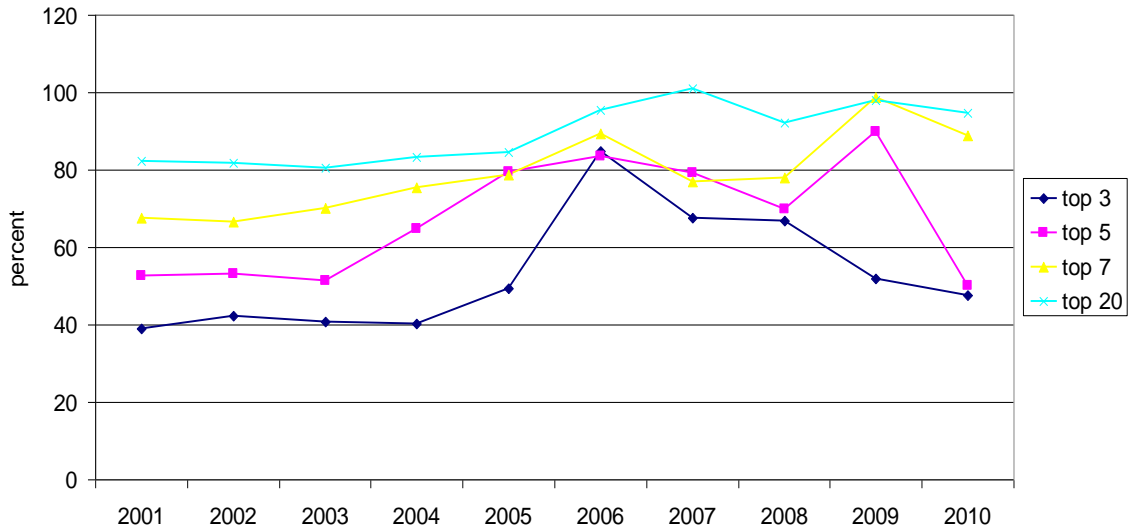
Figure 13: Share of the total capital gain-like revenues of the U.S. BHC sector accruing to the largest BHC.



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

The concentration of the capital gain-like revenues in the largest institutions, however, differs by types of these revenues. For instance, it is the highest for realized capital gains (Figure 14). In 2007-2010 realized capital gains received by the largest 20 banks were equal to capital gains of the BHC sector as a whole. The largest 3 BHC receive about the half of the BHC sector realized capital gains.

Figure 14: Share of realized capital gains of the U.S. BHC sector accruing to the largest BHC.

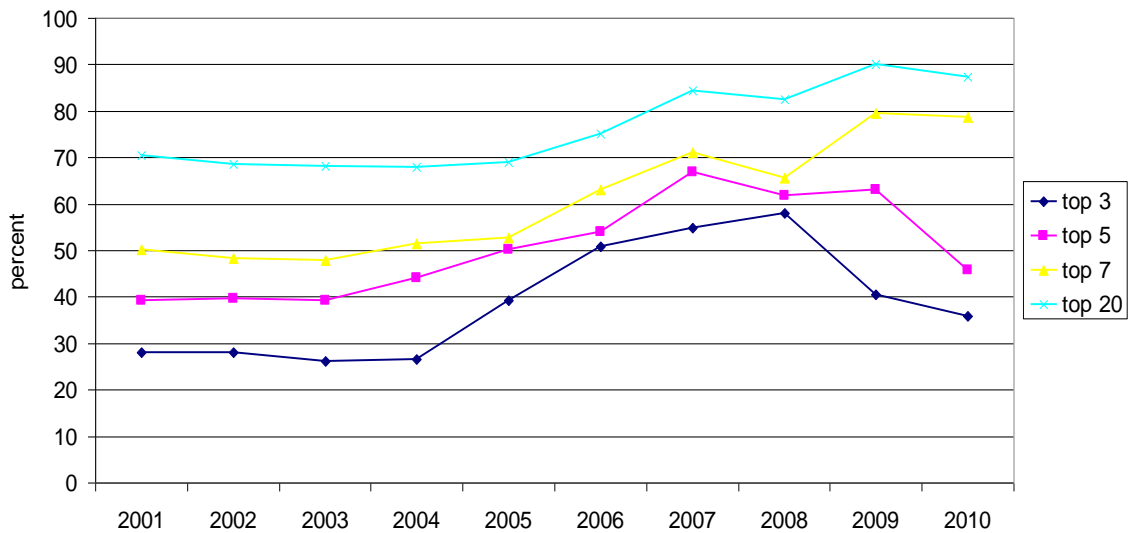


Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

The bank concentration by the size of the founder's profit has also been rising.

The share of the founder's profit received by the largest 20 banks rose from 70 percent in the early 2000s to 90 percent in 2010 (Figure 15).

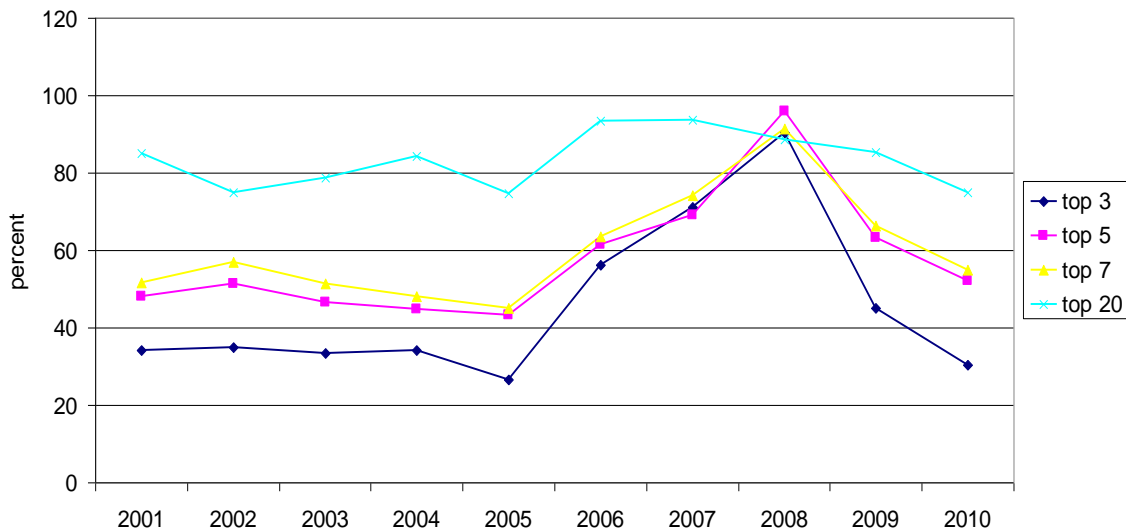
Figure 15: Share of founder's profit of the U.S. BHC sector accruing to the largest BHC.



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

Finally, concentration of BHC by the size of securitization income has been more stable, without a pronounced upward trend (Figure 16). It is also the least concentrated form of capital gain-like revenues, with the largest 20 BHC receiving “only” 80 percent of the total for the sector as a whole. Even this least concentrated form of capital gain-like revenues is highly concentrated.

Figure 16: Share of securitization income of the U.S. BHC sector accruing to the largest BHC.



Source: calculations by author based on FR Y-9C Reports, available through Federal Reserve Bank of Chicago and WRDS.

To recap, an analysis of a rise in BHC noninterest income based on the accounting categories indicates the multiplicity of non-traditional banking activities generating noninterest revenues. This multiplicity conceals, however, a unifying principle behind some of these activities. Namely, several non-traditional banking activities have been driven by extraction of capital gain-like income. This can be seen based on a theoretical argument developed in chapter 5, according to which income due to the gap between the two sets of prices of capital assets takes many forms – underwriting revenue, income from M&A, securitization income, and trading gains. A decomposition of the

BHC noninterest income based on these theoretical categories reveals capital gain-like revenues constituted 40 percent of noninterest income in 2001-2007, on average, and rose to 46 percent in 2009-2010. Securitization income and founder's profit were the largest forms of the capital gain-like revenues before and after the crisis, respectively. Although trading revenues have attracted much attention, they have been the smallest form of capital gain-like revenues prior to the crisis. Even smaller were the "paper profits" associated with the mark-to-market accounting. The capital gain-like revenues are highly concentrated in the largest BHC, with the top 7 banks having received about a half of these revenues in 2001-2005 and 80 percent in 2009-2010. In the same time periods, the largest 20 banks received 75 and 90 percent of total capital gain-like revenues of the BHC sector.

6.4. Beyond the Apparent Heterogeneity of New Banking Activities: Profits from Redistribution of Assets

A decomposition of revenues of the US bank holding companies suggests that almost a half of revenues from non-traditional banking activities have been coming from underwriting, M&A, securitization, and trading in financial assets. Although these activities correspond to different bank functions with respect to other sectors of the economy, and although the concrete mechanisms of revenue extraction differ among them, there is a unifying principle behind these apparently heterogeneous banking businesses. All these activities yield capital gain-like revenues, therefore, extraction of capital gains can be seen as a driving force behind many aspects of the transformation of banking.

The social significance of the capital gain-like revenues being a driving force behind the transformation of banking is rooted in the macroeconomic source of these revenues. As was shown in chapter 5, there are two ways to think about the sources of capital gains, namely, first, as a process creating this type of income, second, as a macroeconomic source related to income and wealth redistribution. It was argued that founder's profit and securitization income come from similar processes. Incorporation involves a change in the nature of capital to truly social capital and, similarly, securitization further raises the social character of banking business, in addition to its inherently social character due to moneyiness of bank liabilities, asset diversification, and an access to reserves. From the perspective of the macroeconomic sources of bank revenues a range of banking activities yielding similar revenues is even wider. It was argued in chapter 5 that bank revenues from underwriting, M&A, securitization, and trading come from the redivision of assets of the intermediate and ultimate asset holders through their accepting a rate of return below the rate of return of the underlying assets. Thus, massive wealth transfers are embedded in functioning of the modern banking, made possible by availability of monetary wealth of the ultimate investors. The greater is the pool of monetary wealth, the greater are the opportunities for extraction of capital gains. Given that the question of sources of capital gain-like revenues as a process was discussed at length in chapter 5 and given that this issue does not require an empirical investigation, for the purposes of an empirical analysis, the rest of the chapter will focus on the other aspect of sources of bank revenues, namely, their macroeconomic sources.

Capital gain-like revenues come from the redivision of assets of others. A rise in wealth accumulated in the form of financial assets compared to the current output,

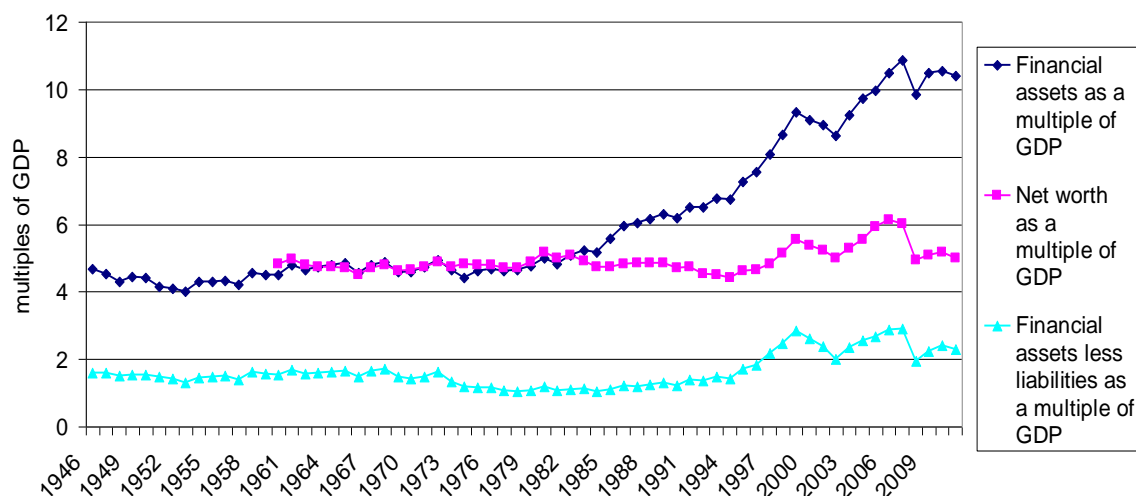
therefore, signifies a rise in the pool of assets that can be redistributed in the form of capital gain-like revenues. Incidentally, Hilferding was aware of this relationship when he argued that founder's profit presupposes the conversion of industrial into fictitious capital, and this convertibility "depends solely upon the quantity of loan capital available which, while retaining the form of interest-bearing capital, is ready to be converted into productive capital. There must be enough money available for investment in shares" (Hilferding 1910, ch. 10). Figure 17 shows that for the US economy net wealth held in the form of financial assets, i.e. financial assets net of liabilities, was declining in 1973-79, then stabilized through the early 1990s, and rose sharply after 1995.¹⁵⁴ Financial assets less liabilities have more than doubled as a multiple of GDP – from 1.15 in 1973-79, on average, to 2.37 in 1995-2011 (Table 4). This rise in the pool of net wealth held in the form of financial assets since the mid-1990s is consistent with a rapid rise in BHC noninterest income as a share of GDP since 1997 (Figure 6). As was shown above, a significant share of the noninterest income represents capital gain-like revenues having a pool of net wealth as its source. Incidentally, the rise in the pool of net wealth held in the form of financial assets is comparable to more than doubling of financial assets as a multiple of GDP since 1946,¹⁵⁵ on the one hand, and stands in contrast to a very modest increase in net worth, on the other hand. Total net worth – a sum of financial and non-

¹⁵⁴ These time periods reflect a regulated social structure of accumulation (SSA) characteristic of the US after the World War II, a crisis of this SSA in 1973-79, and an establishment of a new SSA by the early 1980s (Kotz 2009, 2011).

¹⁵⁵ An increase in financial assets is more dramatic than an increase in financial assets less liabilities due to lengthening of financial intermediation chain.

financial assets less liabilities – as a multiple of GDP has increased by only 11 percent, from the average 4.74 in 1946-1972 to 5.26 in 1995-2011.¹⁵⁶

Figure 17: Total economy financial assets, net worth, and financial assets less liabilities as a multiple of GDP (US, 1946-2011).



Source: calculations by author, Flow of Funds.

Table 4: Total economy financial assets, net worth, and financial assets less liabilities as a multiple of GDP (US, 1946-2011), simple average.

| | Financial assets as a multiple of GDP | Net worth as a multiple of GDP | Financial assets less liabilities as a multiple of GDP |
|-----------|---------------------------------------|--------------------------------|--|
| 1946-72 | 4.53 | 4.74 | 1.54 |
| 1973-79 | 4.63 | 4.78 | 1.15 |
| 1980-94 | 5.87 | 4.80 | 1.23 |
| 1995-2011 | 9.36 | 5.26 | 2.37 |

Source: calculations by author, Flow of Funds.

There are two immediate forms of redistribution of assets, namely, redistribution of the ultimate asset holder's own funds and redistribution of the borrowed funds. At the moment of a financial institution making a capital gain-like profit the macroeconomic source of these revenues differs. In the former case, it is redivision of the already existing

¹⁵⁶ Total economy net worth is calculated as a sum of net worth of nonfinancial corporate business, households and non-profit organizations, financial business, and state, local, and federal government.

wealth (savings), in the latter case, it is a redivision of the newly created assets and liabilities through rising leverage. Given that under normal circumstances a borrower would repay the debt, both cases can be considered as particular instances of a more general case of redistribution of wealth (net worth) of the ultimate asset holder, be it at the moment of transaction or of debt repayment.¹⁵⁷ Transactions with borrowed funds do not alter the fact that the asset holder's net worth forms the ultimate macroeconomic source of the capital gain-like revenues of the financial institutions, they merely change the timing when this redistribution is reflected on the asset holder's balance sheets.

The wealth transfer behind extraction of capital gains has three characteristics. First, the wealth thus redivided is not related to the ultimate asset holder's current income and can therefore exceed it. At the macroeconomic level, this makes capital gain-like revenues incommensurable with revenues reflecting the newly created value in the economy (value added). Accumulated wealth being a source of financial profits independent of the current output is one of the reasons behind the rise in the financial sector profit as a share of total domestic profits, which also explains how this long term trend could be sustained since the onset of the crisis. Although GDP has been sluggish, the stock of accumulated wealth still constitutes a pool that can be redivided in spite of a recession, making the rising share of financial profits in total domestic profits less puzzling.

¹⁵⁷ Debt is not always repaid. For instance, leverage and rolling over debt is immanent to functioning of financial institutions. Moreover, under particular circumstances, economic agents and institutions can default and not repay their debt. In both cases, transactions with borrowed funds while still yielding capital gain-like profits would not lead to an ultimate redivision of assets, therefore, in these cases the assets of others cannot be said to be a source of the capital gain-like revenues.

Second, wealth transfers associated with the capital gain-like revenues are not identical with a zero-sum game. Accepting different rates of return on future cash flows allows both parties to make profit, which in turn justifies the persistence of these gains. This positive-sum game is nevertheless consistent with a redistribution of assets, and the possibility of both parties profiting from a transaction does not make this transaction less of a wealth transfer.¹⁵⁸ It is precisely the apparent win-win game that conceals a wealth transfer accompanying this transaction. Contrary to Haldane, Brennan and Madouros (2010, p. 101), capital gains do not have to be compensated by “paper losses” of those who previously made these profits. The same holds for other capital gain-like revenues. Nor do these revenues have to be matched by losses of the ultimate asset holders, as in a zero-sum game. Nor is it puzzling the capital gain-like revenues do not have to revert to the mean and can instead quickly recover to a pre-crisis level, contrary to Reid’s (2008) expectation that a crisis should wipe out the \$1.2 trillion “excess profits” received by the U.S. financial sector in 1998-2008.¹⁵⁹ This is not to deny that there were structural reasons why a crisis that started in 2008 and losses it brought with it could have been anticipated. For example, Kotz (2009, p. 311) argues that the very structural factors that

¹⁵⁸ The positive sum game nature of the capital gain-like revenues, co-existing with a transfer, is in some sense analogous to a positive sum nature of the capital-labor relationship. Both workers and capitalists receive a net gain in the form of wages and profits, respectively. This net gain for each party co-exists with a transfer of unpaid labor in favor of capital. Unlike the capital-labor relationship that is directly linked to the process of production, the process of extraction of capital gains does not create new value and rather hinges on value created in the sphere of production, but does so indirectly.

¹⁵⁹ Reid (2008) argued that “Given that the WDCI function on Bloomberg reports that \$184bn has been written down by US financials so far in this crisis, if one believes that the size of the financial sector should shrink to levels seen a decade ago then one could come to the conclusion that there is another trillion dollars of value destruction to go in the sector before we're back to the long-run trend in financial profits. A scary thought and one that if correct will lead to a long period of constant intervention from the authorities in an attempt to arrest this potential destruction. Finding the appropriate size of the financial sector in the "new world" will be key to how much profit destruction their needs to be in the sector going forward”. According to Alloway (2010), in 2010 Reid confirmed that a mean reversion of the U.S. financial profits was his expectation in 2008.

underlay an economic expansion of the previous decades – growing income inequality, a speculative financial sector, and a series of asset bubbles – were responsible for a crisis. This crisis resulted in losses for both investors and financial institutions. The fact that profits of the financial institutions were followed by losses across the society holds, however, only for the economy as a whole and says nothing about how these losses were distributed. The losses were born by the ultimate asset holders that might or might not have been the very individuals and institutions that made profits in the previous years. In some cases they were (e.g., AIG, Bear Stearns, Lehman Brothers), in others they were not (e.g., Goldman Sachs). Some of it is also explained by government interventions that bailed out banks and provided other forms of assistance to shield them from potential losses.

Third, capital gain-like revenues involve wealth transfer even in normal times, i.e. even if the value of the asset does not decline. As was shown in chapter 5, accepting differential rates of return is the principle mechanism of such wealth transfers. Asset price deflation, therefore, represents a second-round wealth transfer, in addition to a wealth transfer prior to devaluation. A problem is therefore not only that the asset value collapses in crisis, as has been noticed in the literature on multiple occasions, but rather that transactions in the financial markets are structured in such a way that the ultimate asset holders have assets of indeterminate value, while others lock in profits.

The significance of the wealth transfers associated with capital gains becomes apparent once it is recognized that households hold a substantial share of the total pool of financial assets. Since the World War II, financial assets have amounted to 60-70 percent of household total assets. With a rise of private pensions and other forms of personal

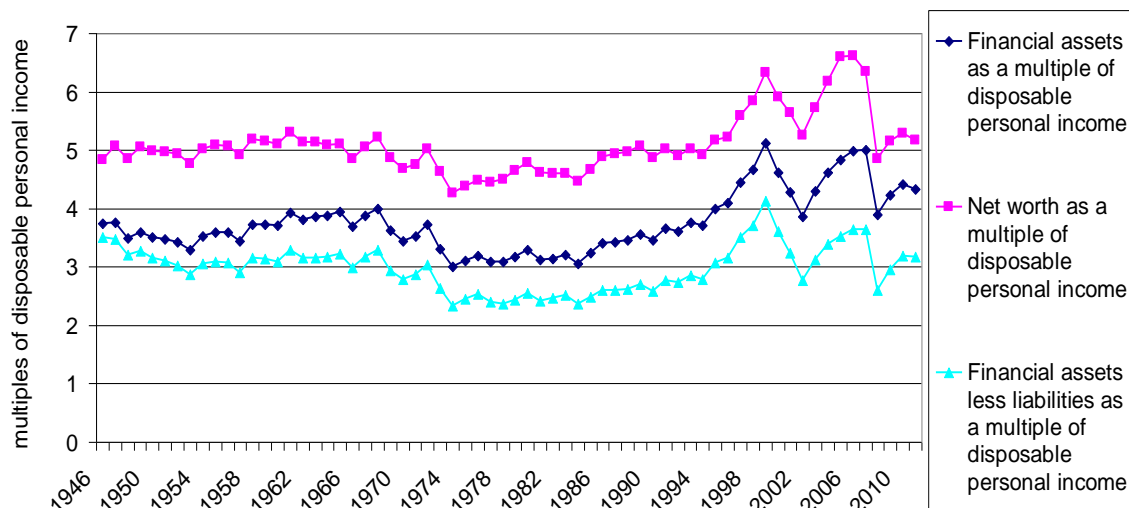
saving, it is household assets and ultimately their net worth that increasingly form a source of the capital-gain like revenues of financial institutions. The wealth transfer happens either directly through redivision of the household saving, or indirectly through increased leverage which would ultimately be followed by redivision of net worth upon the debt repayment.¹⁶⁰ In this context a rise of a stock of household wealth in the form of financial assets and net worth as a multiple of their disposable income acquires a different significance. It signifies not just a rise in household wealth, but also a rise in the stock of funds available to be redivided in the form of capital gain-like revenues of financial institutions.

Figure 18 shows that net wealth held by households in the form of financial assets, i.e. financial assets net of liabilities, has increased by 35 percent between 1973-1979 and 1995-2011 – from the average 2.46 to 3.32 multiples of disposable personal income (Table 5). This figure significantly exceeds a comparable indicator for the economy as a whole – financial assets net of liabilities as a multiple of GDP – which stood at 1.15 and 2.37 in the same time periods (Table 4). It means for the households the pool of net financial assets available for redistribution in the form of capital gain-like revenues forms a larger share of current income than for the economy as a whole, leaving households relatively more exposed to wealth transfers in favor of the financial sector. As

¹⁶⁰ Given that both rolling over debt for purchasing financial assets and defaults by individuals are relatively uncommon compared to businesses, the non-repayment of debt by households can be abstracted from for the present purposes. Households are less likely than other sectors to avoid an ultimate redivision of their wealth through rolling over debt, default, or a bail-out by the government or some other institution. Therefore, their wealth is more likely to form a source of the capital gain-like revenues. There is an in-built asymmetry between households, on the one hand, and financial and non-financial institutions, on the other hand, with respect to the possibilities for their wealth to form an ultimate source of the capital gain-like revenues. A further substantiation of this argument and its empirical illustration lie beyond the scope of this study. For the present purposes suffice it to say that an increase in the household financial assets and in leverage of the financial sector in the course of financialisation can be seen as two sides of the same institutional setting favoring extraction of the capital gain-like revenues.

in the case of the economy as a whole, a sharp (though volatile) rise in net wealth held by households in the form of financial assets as a multiple of their disposable income since the mid-1990s is consistent with a rapid increase in BHC noninterest income as a share of GDP in the same time period (Figure 6).

Figure 18: Households financial assets, net worth, and financial assets less liabilities as a multiple of disposable personal income (US, 1946-2011).



Source: calculations by author, Flow of Funds, Table B.100.

Table 5: Household financial assets, net worth, and financial assets less liabilities as a multiple of disposable personal income (US, 1946-2011), simple average.

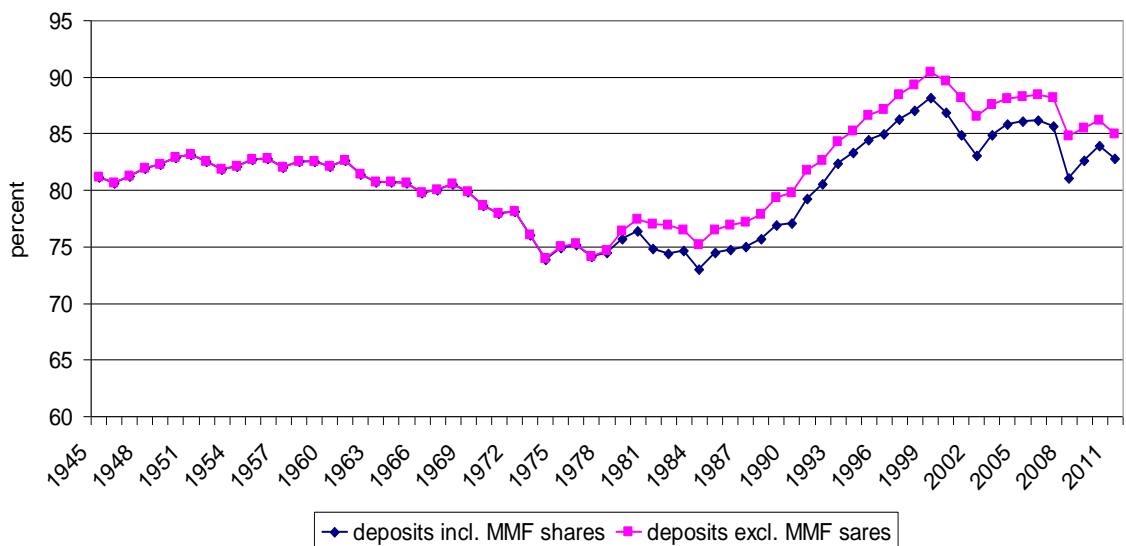
| | Financial assets as a multiple of disposable personal income | Net worth as a multiple of disposable personal income | Financial assets less liabilities as a multiple of disposable personal income |
|-----------|--|---|---|
| 1946-72 | 3.66 | 5.01 | 3.12 |
| 1973-79 | 3.14 | 4.48 | 2.46 |
| 1980-94 | 3.41 | 4.82 | 2.61 |
| 1995-2011 | 4.45 | 5.70 | 3.32 |

Source: calculations by author, Flow of Funds, Table B. 100.

A rising exposure of households to wealth transfers, which is moreover higher than for the economy as a whole, is reinforced by changes in composition of the household assets. Households have been holding a rising share of their assets in forms

other than deposits thus raising the share of the pool of assets available for redivision in the form of capital gain-like revenues. According to Figure 19, household non-deposit financial assets have increased from the average 75 percent of the household total financial assets in 1973-84 to 84 percent in 2000-11. If money market mutual fund shares are excluded from deposits, over the same time period this figure increased even more – from 76 to 87 percent of the household total financial assets. This rise in the non-deposit forms of financial assets means an increase in the share of household assets that can be redivided yielding capital gain-like revenues for financial institutions, hence, a rising vulnerability of households due to a specific form taken by their savings.¹⁶¹

Figure 19: Household non-deposit financial assets as a share of household total financial assets (US, 1945-2011).



Source: calculations by author, Flow of Funds, Table B.100.

At the heart of a rise in household wealth accumulated in a form of non-deposit financial assets net of liabilities, which favors a rise in the capital gain-like income, lie historical and institutional changes associated with a new form of capitalism that emerged

¹⁶¹ An importance of the form taken by savings is also emphasized by Lazonick (1992, p. 473).

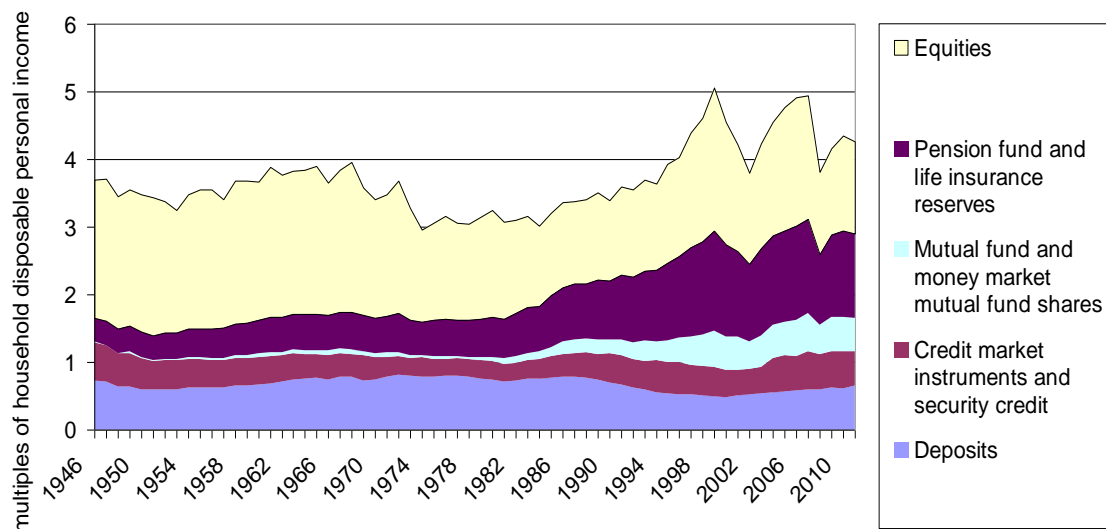
in the US by the early 1980s. This concrete form of capitalism is usually referred to as neoliberalism. A social structure of accumulation approach to capitalist development emphasizes several characteristics of neoliberalism (Kotz 2011, p. 3). Some of these characteristics have directly enabled and sustained a rise to prominence of the capital gain-like revenues.¹⁶²

First, the slashing of state social programs backed by a free market ideology has produced a private pension system, with a rapid rise in the household pension fund reserves. This rise, supplemented by a rise in the household mutual fund and money market fund shares due to a search for returns higher than interest on traditional bank deposits, constituted the main source of a rise in the household financial assets. Privatization of pensions therefore indirectly benefited financial institutions. As shown on Figure 20, pension fund and life insurance reserves have increased from almost a half of household disposable income before 1980 to 81 percent of disposable income in 1980-1994 and 128 percent since 1995. Mutual fund and money market mutual fund shares rose from 4 percent to 18 and 47 percent in the same time periods. A rise in these forms of financial assets added to a traditionally high equity holding by households and resulted in historically high levels of household financial assets. Kotz (2011, p. 15) argued financialisation has its roots in a corporate form of capitalism, with a neoliberal restructuring having removed the barriers that had held in check a proliferation of

¹⁶² A social structure of accumulation (SSA) is “a coherent, long-lasting capitalist institutional structure that promotes profit-making and forms a framework for capital accumulation” (Kotz 2011, p. 2). Neoliberalism can, therefore, be seen as an SSA that raises to prominence a particular form of profit-making, namely, extraction of capital gain-like revenues. Incidentally, as was mentioned in chapter 5, capital gain-like revenues accrue to financial institutions, but also to founders of non-financial enterprises, managers (in the form of stock options), and even high net worth individuals. An extraction of the capital gain-like revenues can therefore be seen as a unifying interest of various capitalist strata helping explain why there was no battle between an identifiable financial and non-financial stratum on the eve of neoliberalism.

financialization. Similarly, a significant size of financial assets held by households – a hallmark of financialization – can be seen as due to household holding of equity shares (reflecting the prevalence of the corporate form of capitalism) being supplemented by pension fund reserves and mutual and money market fund shares enabled by neoliberalism.¹⁶³ Thus, ultimately it is the neoliberal form of capitalism that has allowed for both a dramatic rise in size and a change in composition of the household financial assets that favored a rise to prominence of the capital gain-like revenues.¹⁶⁴

Figure 20: Composition of the household financial assets as a multiple of disposable personal income (US, 1946-2011).



Source: calculations by author, Flow of Funds, Table L.100.

Second, a shift from compromise between labor and capital to a dominance of labor by capital (with aid from the state) manifested in stagnant real wages, even in spite of growing labor productivity. While in 1948-73 output per hour rose at 2.4 percent per

¹⁶³ Neoliberalism has further contributed to an increase in the household financial assets by having been conducive to asset bubbles that caused an increase in the market value of equities held by households.

¹⁶⁴ A need for pension fund reserves in the absence of a publicly provided pension and a shift to money market and mutual fund shares are important causes of “a large and growing volume of investable funds”, in addition to a rise in profit share and a rising concentration of household income discussed by Kotz (2009, p. 308).

year and real average hourly earnings of nonsupervisory workers followed closely growing at 2.2. percent per year, in the neoliberal era the close relationship between labor productivity and remuneration has broken down (Kotz 2009, p. 309). In 1979-2007 output per hour grew at 1.91 percent per year, on average, but average hourly earnings declined at 0.04 percent. The gap between the two rates of growth was wider in 2000-2007 than in the 1990s.

Taken together, these two aspects of neoliberalism, by raising household financial assets and keeping their disposable personal income stagnant, have resulted in a dramatic rise in the household financial assets as a multiple of their disposable personal income (Figure 18). Even a concurrent significant rise in household indebtedness, mostly driven by home mortgages, did not overpower a rise in the household financial assets. As a result, household net financial wealth (financial assets net of liabilities) has also increased as a multiple of their disposable income. This pool of household financial wealth, enabled and sustained by neoliberalism, has created favorable conditions for extraction of capital gain-like revenues and wealth transfers in favor of financial institutions.¹⁶⁵

The empirical findings of the previous section and an appreciation of the capital gain-like revenues being a form of wealth transfer hidden behind the apparently diverse banking activities generating non-interest income offer a lens for evaluating the already existing literature. This literature is right in the emphasis on a problematic character of bank revenues associated with some of the non-traditional banking activities, but this problematic character involves more than just an illusory nature of financial profits or

¹⁶⁵ Neoliberalism has also created favorable conditions for financial profits stemming from problematic accounting practices and increased risk taking. This aspect of financial regulation has attracted a lot of attention in the literature and therefore is not discussed here in detail.

income redistribution through financial mechanisms. Accounting practices and excessive risk taking have played an important role in boosting profit rates in the financial sector and in generating conditions favorable for asset price cycles and a crisis which led to losses. The fact that these profits led to losses in crisis is, however, just one of their attributes which is moreover characteristic of other forms of income. After all, income of all forms declined in a recession. The approach developed here supplements the existing literature by shifting an emphasis to the macroeconomic sources of the financial profit. It is shown that the existence of this profit is not predicated upon future losses. The main contribution of this approach is then twofold. First, it treats the problematic character of revenues from some of the non-traditional banking activities as particular instances of a more general problem – a turn of banking towards extraction of capital gain-like revenues. Second, it is shown that substantial revenues accruing to financial institutions, far from being fictitious, involved real cash flows and systematic wealth transfers behind them. This shift in emphasis results in a new perspective on the social implications of financial profits. The suggested approach supplements the emphasis on mismeasurement of GDP and losses in crisis stressed in the literature by an appreciation of hidden wealth transfers and asset redistribution across the society that, moreover, take place even in normal times.

6.5. Conclusion

This chapter showed that almost a half of revenues associated with non-traditional banking activities represents capital gain-like revenues, making extraction of this form of income a driving force behind many aspects of the transformation of banking. The rising

importance of the capital gain-like income in the total revenues of BHC is non-trivial because it involves wealth transfer from the ultimate asset holders in favor of financial institutions. Neoliberalism enabled and sustained a rise in households' stock of net worth and net financial wealth as a multiple of their disposable personal income, as well as changes in composition of the household financial assets in favor of non-deposit assets. These developments have left households increasingly vulnerable to such wealth transfers. Thus, from the perspective of the macroeconomic sources of bank revenues, the transformation of banking is associated with revenues that are far from being illusory and that involve systematic real wealth transfers in favor of financial institutions.

These wealth transfers have a number of peculiar characteristics. First, being based on differences in the rates of return, they take place even in "normal times" and do not hinge on a depreciation of the asset value in a crisis. Second, these wealth transfers do not have to be a zero-sum game, as differences in the rates of return allow all parties to profit while wealth is still being redistributed. This positive-sum game nature of the gains gives a social justification to the wealth transfers and makes them sustainable. Third, coming from a stock of accumulated assets or being based on a rising leverage in anticipation of a future stock of assets, capital gain-like revenues are unrelated to and hence incommensurable with the current output. This incommensurability makes a rise in the financial profits as a share of total domestic profits less surprising, but with that not less problematic from the perspective of wealth transfers and their social implications.

CHAPTER 7

CONCLUSION

The present study makes four contributions. First, it compares a Marxist theory of banking with other theoretical approaches to banking that emerged to explain its ongoing transformation. A Marxist theory is shown to retain strengths and overcome weaknesses of other takes on banking. This comparative advantage stems from a dual foundation of this theory – liquidity provision through exchange of promises to pay as a core function of banking and an explicit connection between bank revenues and key macroeconomic aggregates (profits, wages, and assets of various sectors). Second, some of the characteristics of founder's profit are shown to hold for profits from securitization and to some extent gains from mergers and acquisitions. Thus, this study identifies a range of revenues associated with functioning of capital markets – capital gain-like revenues – that differ from profits from production, as they involve transfers of wealth. Third, due to their function of liquidity provision, banks are well suited to facilitate the deals yielding capital gain-like revenues and, hence, to share in these gains. It implies a relatively unaltered fundamental function of banking can not only co-exist with but also form a basis for a significant transformation of the nature of bank revenues. Moreover, extraction of the capital gain-like revenues is a unifying principle behind a number of apparently heterogeneous aspects of the transformation of banking. Finally, fourth, an empirical analysis for the US bank holding companies shows capital-gain like revenues were indeed a significant part of bank revenues in 2001-2010. In the context of rising household vulnerability to wealth transfers, such trend in bank income suggests a

reinstatement of predatory aspects of finance on a new foundation. Consider these arguments in a greater detail.

7.1. Comparative Advantage of a Marxist Theory of Banking

Since the 1990s new approaches to banking have emerged trying to fill a theoretical vacuum that arose due to weaknesses of the information-theoretic approach in addressing the ongoing transformation of banking. The present study compares a Marxist theory of banking with other approaches to banking, both mainstream and heterodox, and argues a Marxist framework is better suited for understanding the ongoing transformation of banking. It is shown to retain strengths and overcome limitations of other approaches to banking.

A comparative advantage of a Marxist theory in explaining the transformation of banking stems from its dual foundation (chapter 4). First, it is grounded in the bills view, according to which banks originate in discounting, and the essence of the banking business lies in issuing one promise to pay (bank notes) in exchange for another (discounted bills). Second, Marx distinguishes between different types of credit relationship based on characteristics of a borrower and a lender, and on a historical context in which they take place. This offers a useful framework for locating banking in the circuit of capital and revenue by establishing a connection between various forms of bank profit and their sources. For Marx, profit in the sphere of circulation can come not only from redivision of the newly created surplus value, as in the case of merchant's profit, profit from money-dealing and interest from lending to industrial capitalists. Drawing on Stuart (1770a, p. 206), Marx argues profit in the sphere of circulation can

also come from redivision of the already existing value – wages and existing money stocks. In this case it would be “profit upon alienation”, which involves “a vibration of the balance of wealth between parties” – a transfer of income or wealth.

A Marxist theory of banking can be seen as a more general approach to banking than the mainstream or the circuitist theories of banking. The former retains a trans-historical dimension of the banking business while still allowing for a significant transformation of banking associated with a change in the social nature and macroeconomic sources of bank revenues and profits.

First, a Marxist theory of banking retains the strength of the new mainstream and heterodox approaches to banking by identifying a core of the banking business outside of the context of information asymmetry and market imperfections. It allows these theories to overcome the inability of the information-theoretic approach to account for a rising share of bank profits in total domestic profits (chapter 2). A Marxist theory does it by defining banking as liquidity provision through an exchange of promises to pay, hence, makes banks a specifically monetary phenomenon. It still leaves room for screening and monitoring functions of banking, but merely as necessary aspects of banking business required for liquidity provision, not as *raison d'être* of banks. Banks would have existed even if information were symmetrical, as long as money is a universal equivalent and demand for universal equivalent poses a need for credit relations. Similarly, banks can receive a rising share of domestic profits with less screening and monitoring if they find a new form of liquidity provision.

Second, a Marxist theory of banking can nest the portfolio theories (Fama 1980, Gurley, Shaw 1956, Tobin 1963) and more recent theories of shadow banking as banking

(Gorton 2008, 2009, Pozsar et al. 2010) as its particular cases (chapter 2). It becomes possible because, by focusing on a transformation of one class of assets into another, the portfolio theories and theories of shadow banking can be seen as partial cases of the bills view tailored for the realities of banking in the middle of the 20th and early 21st centuries, respectively.¹⁶⁶ Grounded in the bills view, a Marxist theory of banking defines banking as a business of exchange of promises to pay in general, allowing promises to pay and forms of their exchange take many concrete forms, from discounting through credit intermediation to the modern forms of shadow banking and asset management. In addition to focusing only on particular historical forms of banking, all the recent mainstream approaches to banking overlook the transformation of banking associated with a changing character of bank profits. This aspect of the transformation can be recognized by a Marxist theory which, unlike the mainstream theories, locates banking in the macroeconomic context and basic macroeconomic aggregates, such as various forms of income and stocks of monetary wealth.

Third, a Marxist theory retains Keynes's emphasis on the role of liquidity (chapter 3), because an exchange of promises to pay as a trans-historical function of banking has liquidity provision as its counterpart. A credit relationship is an exchange of money for a promise to pay. This relationship therefore involves two aspects – a monetary aspect of liquidity provision and an issuance of a promise to pay – each corresponding to a different function of money, namely, money as a universal equivalent and money as a means of payment. By supplementing the focus on liquidity provision by an emphasis on exchange of promises to pay, a Marxist theory lowers the level of abstraction at which

¹⁶⁶ Similarly, banking for development (Schumpeter 1961 [1911/1934]) is but a particular form of banking relevant under certain historical circumstances without being a general theory of banking (chapter 3).

banking is theorized – from a high level of abstraction of a monetary theory to a more concrete level of credit relations. This change in the level of abstraction brings to the surface the role of trust, power, and other non-economic factors.

Fourth, an explicit connection between banking and the circuit of capital and revenue, as well as a take on banking as a monetary phenomenon, makes a Marxist theory akin to an approach to banking within the monetary theory of production (chapter 3). Nevertheless, the monetary circuit of production theory cannot explain several aspects of the transformation of banking. First, if the core function of the banking business is money creation *ex nihilo*, it is unclear why banks take deposits, and why they engage in asset sales and securitization. Moreover, this approach excludes the shadow banking system from the perimeter of banking due to a lack of money creation by the shadow banking institutions, thus, leaving out of the scope of analysis an important part of the modern financial system. Second, a rise in lending to households and non-lending activities reveal the limitations of centering a theory of banking on credit for productive purposes, as is the case with the circuitist theory of banking. Assuming only one form of connection between banking and the rest of the economy leads circuitists to explain lending for non-productive purposes as a roundabout way of extending productive credit. This attempt to preserve an analytical core of the theory prevents the circuitist approach from realizing a full potential of an explicit connection between banking and the circuit of capital, embedded in the theory. Third, the rising multiplicity of forms of bank revenues that moreover have different sources stand at odds with the circuitist treatment of bank profit as a homogeneous rent, or seigniorage, extracted from the rest of the economy. Therefore, a Marxist theory appears to be a more general theory of banking

than the monetary theory of production, because the former views money creation ex nihilo as one of the forms of exchange of promises to pay and because it allows for multiple links between banking and the circuits of capital and revenue, not restricted to credit for productive purposes.

7.2. A Range of Capital Gain-Like Revenues

The second contribution of the present study lies in identifying a range of capital gain-like revenues having similar characteristics with respect to the social relations making their extraction possible and the macroeconomic source of these gains. Certain features of founder's profit are shown to hold for profit from securitization and to some extent gains from mergers and acquisitions.

A genesis of the corporate form of business organization and capital markets has become a social foundation for a particular form of profit making – extraction of capital gain-like revenues (chapter 5). An analytical tool for understanding the peculiar character of these revenues can be found in a distinct conceptualization of capital goods as having two sets of prices – the market price based on capitalization of their future yields and the replacement costs. An insight that a gap between the two sets of prices can be seen as a gain motivating investment is a basis of a theory of investment associated with Keynes (1973 [1936]), Brainard and Tobin (1968, 1969, 1990), and Minsky (2008 [1986], 1975, 1981). This gap would be however an inducement to invest only under particular circumstances, namely, when a decision maker about investment is also remunerated by the gain associated with the gap between the two sets of prices, or at least by a share in this gain. It is therefore not surprising this theory does not stand empirical evidence as a

general theory of investment. In spite of a limited explanatory power of the Keynes-Tobin-Minsky theory of investment, its significance lies in uncovering a peculiar source of profit associated with two ways of pricing the same capital good (forward-looking and “present-looking”). This idea forms a point of departure for the main contribution of the present study – an identification of a range of profits having similar characteristics. Capital gain-like revenues are shown to be a driving force behind IPO, M&A, and securitization.

The simplest form of profit due to a gap between the two sets of prices is founder’s profit – a gain that arises upon conversion of an individually owned enterprise into a corporation. An identification of this profit as an economic category *sui generis* is Hilferding’s (1910) theoretical contribution. Founder’s profit emerges due to a lack of control over enterprise by shareholders, which is translated into their accepting a rate of return below the profit rate. This gain is further magnified by information asymmetry and fraud for which the corporate form of business organization creates a favorable environment. The social significance of founder’s profit stems from its coming from monetary assets scattered across the society, either immediately (if the shares are purchased with the ultimate investor’s own funds, or savings) or after a period of time (upon repayment of debt, if the shares are purchased with borrowed funds). Capital gain-like revenues can therefore be seen as a form of wealth transfer, or “profit upon alienation”, the scope of which is unrelated to and can exceed the size of the current income.

Mergers and acquisitions is another way of extracting capital gain-like revenues. Unlike founder’s profit, a gain from M&A arises due to an increase in the expected future

profit of a combined enterprise, not due to a difference in the required rates of return. In spite of this difference, profit from M&A is akin to founder's profit in the sense of coming from an increase in the current valuation of a reorganized enterprise, which opens an opportunity for a lump-sum gain. This gain accrues to institutions and individuals having control over the enterprises under reorganization and to institutions facilitating this process (e.g., investment banks). Coming from the monetary assets of the buyers of securities of a reorganized enterprise, gain from M&A also constitutes a form of "profit upon alienation".

Finally, securitization is akin to incorporation both in terms of the process and social foundations of the financial deal, and the character of the gain. Securitization can be seen as a process of incorporation applied to bank lending, which increases the current valuation of a loan. Consider the similarities. First, similar to IPO involving a transformation of individual capital into truly social by dispersing the ownership structure, securitization raises the social foundations of bank credit to a new level. All bank credit generalizes trust and social foundations of the trade credit by overcoming the limits of credit confined to intrinsically linked production processes, requiring a match in volume and maturity, and lacking universal acceptability (Itoh, Lapavitsas 1999, p. 90-91, Lapavitsas 2003, p. 79-81, 2007, p. 426-427). Banks broaden the basis of trust generalizing credit relations beyond the confines of related industrial capitalists through asset diversification and access to reserves, thus, giving bank liabilities a character of a generally accepted universal equivalent. The specificity of securitization lies in its generalizing the social foundations of credit without bearing the costs of reserves and consequently also lowering the costs of screening and monitoring. Securitization can

therefore be seen as a way of generalizing credit relations at a minimum cost for banks. Second, gains from securitization come from a lack of control of the ultimate asset holders over banks, resulting in their accepting a rate of return below the interest rate on loans. Third, similar to IPO, gains from securitization are often further magnified by fraud and information asymmetry for which the difference in the social status of insiders and outsiders creates a favorable environment. Finally, fourth, both cases involve a transformation of an illiquid asset into a liquid one – means of production into equity shares and loans into marketable securities (bonds). Although this difference in liquidity of the assets can explain some of the difference in the required rates of return from the perspective of an individual investor, it overlooks a conflict between the social nature of liquidity and private appropriation of its benefits by insiders and financial institutions. If incorporation or securitization were handled by a public agency, the benefits of liquidity – social in nature – would have belonged to the general public.

There are three important characteristics of the capital gain-like revenues (chapter 6). First, being based on differences in the rates of return, the wealth transfers associated with the capital gain-like revenues take place even in “normal times” and do not hinge on a depreciation of the asset value in a crisis. Second, these wealth transfers do not have to be a zero-sum game, as differences in the rates of return allow all parties to profit while wealth is still being redistributed. This positive-sum game nature of the gains gives a social justification to the wealth transfers and makes them sustainable. Third, coming from a stock of accumulated assets or being based on a rising leverage in anticipation of a future stock of assets, capital gain-like revenues are unrelated to and, hence, incommensurable with the current output. This incommensurability makes a rise in

financial profits as a share of total domestic profits less surprising, but with that not less problematic from the perspective of wealth transfers and their social implications.

7.3. Extraction of Capital Gain-Like Revenues as a Driving Force behind Many Aspects of the Transformation of Banking

This study contributes to the existing literature by arguing that some of bank profits represent capital gain-like revenues and that extraction of this form of revenues has been moreover a unifying principle behind many aspects of the ongoing transformation of banking. This conclusion can be reached by integrating a Marxist theory of banking with a suggested take on the capital gain-like revenues associated with the dual price system. The former emphasizes the role of banks in liquidity provision through exchange of promises to pay, which serves as a foundation for banks' sharing in the capital gain-like revenues. The latter stresses multiple forms taken by the capital gain-like revenues and uncovers their peculiar character.

A dual foundation of a Marxist theory of banking makes this theory well suited to explain how a relatively unchanged core function of banking can not only co-exist with but also form a basis for a significant transformation of bank revenues. IPO, M&A, securitization and trade in assets rose to prominence with a rise of the corporate form of business organization and capital markets (chapter 5). These activities yield a peculiar form of profit, capital gain-like revenue, but they also require liquidity. Banks have responded to a rise in these activities by deploying their function of liquidity provision through an exchange of promises to pay. On these grounds, banks share in the capital-gain like revenues which in turn potentially makes them active promoters of these forms

of profit making. This is in line with an active principle behind the banking business in general stressed by a Marxist theory of banking.

An appreciation of the multiplicity of forms taken by the capital gain-like revenues allows one to recognize that behind an apparent heterogeneity of banking activities associated with the transformation of banking lies a common driving force – extraction of the capital gain-like revenues (chapter 6). This unifying principle would not be noticed, if an analysis were based on the accounting categories as given in bank income statements.

How does this analysis of the transformation of bank revenues relate to other approaches? An emphasis on bank revenues is in line with the literature on financialization (Epstein, Jayadev 2005, Krippner 2005) stressing a rise in financial profits (chapter 3) and the mainstream empirical literature (Samolyk 2004, DeYoung, Rice 2004b) treating bank revenues and profits as an appropriate lens for studying the transformation of banking due to a peculiar character of this transformation that is not always reflected on bank balance sheets (chapter 2). Nevertheless, by identifying a range of capital gain-like revenues, the present study goes beyond an empiricist bias of the mainstream literature focusing on the forms of revenues as given in the bank income statements. In this sense the present approach draws on the literature on financialization (Pollin 1996, van Treeck 2009, Lapavitsas 2009, Dos Santos 2009) posing a question of the macroeconomic sources of financial profits (chapter 3). An appreciation of a variety of the macroeconomic sources of financial profits allows this approach to go beyond the circuitist theory, which treats bank profits as a homogeneous rent. An identification of accumulated stock of monetary wealth as a distinct macroeconomic source of financial

profit contributes to Arrighi-Pollin debate (Pollin 1996, Arrighi 1994, 1997) by supporting Pollin's claim that financial profits ultimately come from production. Nevertheless, an approach developed in the present study shows the complexity of the relationship between production and financial profit. In the case of the capital gain-like revenues, financial profit comes from an accumulated stock of wealth that is unrelated to current output, hence, to the current process of production. Finally, the focus on the macroeconomic sources of bank revenues supplements the existing studies focusing on the fictitious character of these revenues, due to either accounting standards or risk illusion (Kerr 2011, Crotty 2008, Haldane, Brennan & Madouros 2010) (chapter 6). Although capital gain-like revenues can be considered fictitious in the sense that their magnitude can change suddenly for reasons unrelated to the process of production, from the perspective of their macroeconomic sources, far from being illusory, these profits involve a real wealth transfer. By identifying a structural problem with extraction of some forms of financial profit, this study suggests that addressing the issues of accounting practices and excessive risk taking is not sufficient to eliminate the capital gain-like revenues.

7.4. Significance of the Capital Gain-Like Revenues for the US Bank Holding Companies in 2001-2010

A theoretical framework developed in the present study is used to empirically examine the transformation of banking in the US. An empirical analysis of the sources of revenue of the US bank holding companies in 1986-2010 shows the capital gain-like revenues have indeed been a significant part of bank income (chapter 6). In 2001-2010,

almost a half of the non-interest income of the US bank holding companies came from activities generating capital gain-like revenues, with the share of this income in 2009-2010 exceeding its pre-crisis level. These revenues are highly concentrated in a small number of the largest banks. In 2009-2010, the largest seven banks received 80 percent, and largest twenty banks – 90 percent of the total capital gain-like revenues accruing to the bank holding companies sector as a whole.

Neoliberalism has enabled and sustained a rise in the household financial assets as a multiple of their disposable income through slashing state social programs, including privatization of pensions, and keeping the real wages stagnant. Household net wealth (financial assets net of liabilities) as a multiple of their income has increased by 35 percent between 1973-1979 and 1995-2011 – from the average 2.46 to 3.32. This rise was reinforced by a changing composition of the household financial assets. Household non-deposit financial assets (including money market mutual fund shares) have increased from the average 76 percent of the household total financial assets in 1973-84 to 87 percent in 2000-11. These two developments indicate a rise in the pool of assets available for redivision in the form of capital gain-like revenues, and suggest households have become increasingly vulnerable to wealth transfers associated with extraction of the capital gain-like revenues. Financial assets represent a larger multiple of the household disposable income than a comparable indicator for the US economy as a whole (1.15 in 1973-1979 and 2.37 in 1995-2011) revealing a greater exposure of households to wealth transfers than other sectors. This high exposure is reinforced by households being less likely to avoid the ultimate wealth transfers through defaults, rolling over debt, or bail-outs by the government or other institutions.

In the context of rising household exposure to wealth transfers, a rise in the capital gain-like revenues of the bank holding companies signifies a reinstatement of predatory aspects of finance in contemporary capitalism.

BIBLIOGRAPHY

Data sources:

1. The NIPA data are available at http://bea.gov/iTable/index_nipa.cfm, accessed on 06/20/12, last revised on 05/31/12.
2. FDIC data are available at <http://www2.fdic.gov/hsob/>, accessed on 06/20/12, last updated in March 2012.
3. The FR Y-9C reports are available from Research Data Services (WRDS), The Bank Regulatory Database, at <http://wrds-web.wharton.upenn.edu/wrds/index.cfm?true>, accessed on 05/27/12.
4. The Flow of Funds data are available at <http://www.federalreserve.gov/datadownload/Build.aspx?rel=Z1>, accessed on 06/19/12, last updated on 06/07/12.

References:

- Adrian, T. & Shin, H.S. 2010, "Liquidity and Leverage", *Journal of Financial Intermediation*, vol. 19, no. 3, pp. 418-437.
- Allen, F. & Santomero, A.M. 2001, "What do financial intermediaries do?", *Journal of banking finance*, vol. 25, no. 2, pp. 271.
- Allen, F. & Santomero, A.M. 1997, "The theory of financial intermediation", *Journal of Banking & Finance*, vol. 21, no. 11-12, pp. 1461-1485.
- Alloway, T. 2010, *The Bank Problem in a Single Chart*, Financial Times.
- Arrighi, G. 1994, *The Long Twentieth Century: Money, Power, and the Origins of our Times*, Verso, London, New York.
- Arrighi, G. 1997, "Financial Expansions in World Historical Perspective: A Reply to Robert Pollin", *New Left Review*, vol. I, no. 224, pp. 154-159.
- Ashcraft, A.B. & Steindel, C. 2008, "Measuring the Impact of Securitization on Imputed Bank Output", *Federal Reserve Bank of New York*.
- Bagehot, W. 1877, *Lombard Street: A Description of the Money Market*, Scribner, Armstrong & Co, New York.
- Baker, D. 2011, *False Profits: Recovering from the Bubble Economy*, Berrett-Koehler Publishers, Williston, VT, USA.

- Basu, S., Inklaar, R. & Wang, J.C. 2008, *The Value of Risk: Measuring the Service Output of U.S. Commercial Banks*, Federal Reserve Bank of Boston.
- Berle, A.A. & Means, G.C. 1932, *The Modern Corporation and Private Property*, Harcourt, Brace & World, Inc., New York.
- Berlin, M. & Mester, L.J. 1998, "Deposits and Relationship Lending", *Federal Reserve Bank of Philadelphia Working Paper no. 96-18/R*.
- Bernanke, B.S. 2008, "Addressing Weaknesses in the Global Financial Markets: The Report of the President's Working Group on Financial Markets", *Speech at the World Affairs Council of Greater Richmond's Virginia Global Ambassador Award Luncheon, Richmond, Virginia, April 10, 2008*.
- Bezemer, D.J. 2012, *Finance and Growth: When Credit Helps, and When It Hinders*, Paper Presented at INET Conference, Berlin, April 2012.
- Bhattacharya, A.K., Berliner, W.S. & Fabozzi, F.J. 2008, "The Interaction of MBS Markets and Primary Mortgage Rates", *The Journal of Structured Finance*, vol. 14, no. 3, pp. 16-36.
- Blaug, M. 1996 [1962], *Economic Theory in Retrospect*, Cambridge University Press, Cambridge.
- Bodnaruk, A., Massa, M. & Simonov, A. 2009, "Investment Banks as Insiders and the Market for Corporate Control", *Review of Financial Studies*, vol. 22, no. 12, pp. 4989-5026.
- Bolbol, A.A. & Lovewell, M.A. 2001, "Three Views on Stock Markets and Corporate Behavior: Tobin, Veblen, and Marx", *Journal of Post Keynesian Economics*, vol. 23, no. 3, pp. 527-543.
- Borio, C. 2010, "Ten Propositions about Liquidity Crises", *CESifo Economic Studies*, vol. 56, no. 1, pp. 70-95.
- Borio, C. 2007, "Change and Constancy in the Financial System: Implications for Financial Distress and Policy", *BIS Working Paper No. 237*.
- Bossone, B. 2003, "Thinking of the Economy as a Circuit" in *Modern theories of money: the nature and role of money in capitalist economies*, eds. L. Rochon & S. Rossi, Edward Elgar, Cheltenham, Northampton, pp. 142-172.
- Bossone, B. 2001a, "Circuit theory of banking and finance", *Journal of Banking & Finance*, vol. 25, no. 5, pp. 857-890.

- Bossone, B. 2001b, "Do banks have a future?: A study on banking and finance as we move into the third millennium", *Journal of Banking & Finance*, vol. 25, no. 12, pp. 2239-2276.
- Bossone, B. 2000, "What Makes Banks Special? A Study of Banking, Finance, and Economic Development", *The World Bank Policy Research Working Paper No. 2408*.
- Bosworth, B. 1981, "Taxation and Corporation Investment: A Q-Theory Approach. Comment on L. H. Summers. ", *Brookings Papers on Economic Activity*, vol. 1, pp. 128-132.
- Boyd, J.H. & Gertler, M. 1994, "Are Banks Dead? Or Are the Reports Greatly Exaggerated?", *Federal Reserve Bank of Minneapolis Quarterly Review*, vol. 18, no. 3.
- Boyd, J.H. & Gertler, M. 1993, "U.S. Commercial Banking: Trends, Cycles, and Policy", *NBER Macroeconomics Annual*, vol. 8, pp. pp. 319-368.
- Brainard, W.C. & Tobin, J. 1968, "Pitfalls in Financial Model Building", *The American Economic Review*, vol. 58, no. 2, Papers and Proceedings of the Eightieth Annual Meeting of the American Economic Association, pp. 99-122.
- Braudel, F. 1983, *The Wheels of Commerce*, William Collins Sons & Co Ltd, London.
- Brunnermeier, M.K. & Pedersen, L.H. 2009, "Market Liquidity and Funding Liquidity", *Review of Financial Studies*, vol. 22, no. 6, pp. 2201-2238.
- Campbell, M. 2002, "The Credit System" in *The Culmination of Capital: Essays on Volume III of Marx's Capital*, eds. M. Campbell & G. Reuten, Palgrave, London, New York.
- Cantillon, R. 2001 [1755/1931], *Essay on the Nature of Commerce in General* Transaction Publishers, New Jersey.
- CGFS 2011, *Global Liquidity – Concept, Measurement and Policy Implications*, CGFS Papers No 45, Committee on the Global Financial System, BIS, Basel.
- Chandler, A.D. 1977, *The Visible Hand: The Managerial Revolution in American Business*, Harvard University Press, Cambridge, London.
- Chick, V. 1992, *On Money, Method, and Keynes: Selected Essays*, Macmillan, London.
- Clark, P. 1979, "Investment in the 1970s: Theory, Performance, and Prediction", *Brookings Papers on Economic Activity*, vol. 1, pp. 73-124.
- Colangelo, A. & Inklaar, R. 2010, *Banking Sector Output Measurement in the Euro Area – A Modified Approach*, University of Groningen.

- Corrigan, E.G. 2000, "Are banks special? A revisitation", *The Region*, vol. 14, no. 1, pp. 14-17.
- Corrigan, E.G. 1982, "Are banks special?", *Annual report*.
- Crotty, J. 2008, "If Financial Market Competition is Intense, Why are Financial Firm Profits so High? Reflections on the Current 'Golden Age' of Finance", *Competition & Change*, vol. 12, no. 2, pp. 167-183.
- Crotty, J. 1990, "Owner-Manager Conflict and Financial Theories of Investment Instability: A Critical Assessment of Keynes, Tobin, and Minsky", *Journal of Post Keynesian Economics*, vol. 12, no. 4, pp. 519-542.
- Cutler, A., Hindess, B., Hirst, P. & Hussain, A. 1977, *Marx's 'Capital' and Capitalism Today. Vol. 2*, Routledge & Kegan Paul, London, Henley and Boston.
- DeYoung, R. & Rice, T. 2004a, "How do banks make money? The fallacies of fee income", *Economic perspectives*, vol. 28, no. 4, pp. 34-51.
- DeYoung, R. & Rice, T. 2004b, "Noninterest Income and Financial Performance at US Commercial Banks", *Federal Reserve Bank of Chicago*.
- Diamond, D.W. 1984, "Financial Intermediation and Delegated Monitoring", *The Review of Economic Studies*, vol. 51, no. 3, pp. 393-414.
- Diamond, D.W. & Dybvig, P.H. 1983, "Bank Runs, Deposit Insurance, and Liquidity", *The Journal of Political Economy*, vol. 91, no. 3, pp. 401-419.
- Diamond, D.W. & Rajan, R.G. 2001a, "Banks and Liquidity", *The American Economic Review*, vol. 91, no. 2, pp. 422-425.
- Diamond, D.W. & Rajan, R.G. 2001b, "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking", *The Journal of Political Economy*, vol. 109, no. 2, pp. 287-327.
- Diamond, D.W. & Rajan, R.G. 2000, "A Theory of Bank Capital", *The Journal of Finance*, vol. 55, no. 6, pp. 2431-2465.
- DiNapoli, T.P. 2008, *The Securities Industry in New York City, Report 7-2009*, New York State Comptroller, available at <http://www.osc.state.ny.us/osdc/rpt7-2009.pdf>.
- Dos Santos, P. 2009, "On the Content of Banking in Contemporary Capitalism", *Historical Materialism*, vol. 17, no. 2, pp. 180-213.
- Edwards, F.R. 1993, "Financial Markets in Transition -- or the Decline of Commercial Banking", *Proceedings*, pp. 5-69.

- Edwards, F.R. & Mishkin, F.S. 1995, "The Decline of Traditional Banking: Implications for Financial Stability and Regulatory Policy", *Economic Policy Review*, vol. 1, no. 2, pp. 27-45.
- Epstein, G.A. & Jayadev, A. 2005, "The Rise of Rentier Incomes in OECD Countries: Financialization, Central Bank Policy and Labor Solidarity" in *Financialization and the World Economy*, ed. G.A. Epstein, Edward Elgar, London.
- Epstein, G. & Power, D. 2003, "Rentier Incomes and Financial Crises: An Empirical Examination of Trends and Cycles in Some OECD Countries", *Canadian Journal of Development Studies*, vol. 29, no. 2, pp. 230-248.
- Ertürk, I. 2009, "Banks Working for Themselves: Financial Intermediation and the Banking Crisis of 2007", *Rio's International Journal on Sciences of Industrial and Systems Engineering and Management*, vol. 2.
- Fabozzi, F.J. (ed) 2001, *The Handbook of Mortgage-Backed Securities*, McGraw-Hill.
- Fama, E.F. 1980, "Banking in the Theory of Finance", *Journal of Monetary Economics*, vol. 6, no. 1, pp. 39-57.
- Feldman, R. & Lueck, M. 2007, "Are Banks Really Dying this Time?", *The Region*, vol. 21, no. 3, pp. 6-51.
- Fine, B. 2010, "Locating Financialisation", *Historical Materialism*, vol. 18, no. 2, pp. 97-116.
- Fisher, S. & Merton, R. 1984, *Macroeconomics and Finance: The Role of the Stock Market*, National Bureau of Economic Research, Working Paper No. 1291.
- Friedman, B.M. 2009, "The Failure Of The Economy And The Economists", *New York Review of Books*, vol. LVI, no. 9, pp. 42-45.
- Gertler, M. 1988, "Financial structure and aggregate economic activity: an overview", *Journal of Money, Credit, and Banking*, vol. 20, no. 3, pp. 559-588.
- Gorton, G. 2009, "Slapped in the Face by the Invisible Hand: Banking and the Panic of 2007", *Federal Reserve Bank of Atlanta's 2009 Financial Markets Conference: Financial Innovation and Crisis*.
- Gorton, G.B. 2008, "The Panic of 2007", *NBER Working Paper 14358*.
- Gorton, G. 1994, "Bank Regulation When 'Banks' and 'Banking' Are Not the Same", *Oxford Review of Economic Policy*, vol. 10, no. 4, pp. 106.

- Gorton, G. & Pennacchi, G. 1990, "Financial Intermediaries and Liquidity Creation", *Journal of Finance*, vol. 45, no. 1, pp. 49-71.
- Graziani, A. 2003, *The Monetary Theory of Production*, Cambridge University Press, Cambridge, New York.
- Graziani, A. 1987, "Keynes' Finance Motive", *Economies et Societes*, vol. 21, no. 9, pp. 23-42.
- Greenspan, A. 1994, "Optimal Bank Supervision in a Changing World", *Remarks before the 30th Annual Conference on Bank Structure and Competition*.
- Gurley, J.G. & Shaw, E.S. 1956, "Financial Intermediaries and the Saving-Investment Process", *The Journal of Finance*, vol. 11, no. 2, pp. pp. 257-276.
- Hakenes, H. 2004, "Banks as delegated risk managers", *Journal of Banking & Finance*, vol. 28, no. 10, pp. 2399-2426.
- Haldane, A., Brennan, S. & Madouros, V. 2010, "What is the contribution of the Financial Sector: Miracle or Mirage?" in *The Future of Finance: The LSE Report* London School of Economics and Political Science, pp. 87-120.
- Harris, L. 1976, "On Interest, Credit and Capital", *Economy and Society*, vol. 5, no. 2, pp. 145-177.
- Harvey, D. 2006, *The Limits to Capital*, Verso, London, New York.
- Herring, R.J. & Santomero, A.M. 2000, "What Is Optimal Financial Regulation?", *Center for Financial Institutions Working Paper 00-34*.
- Hicks, J. 1989, *A Market Theory of Money*, Clarendon Press, Oxford.
- Hicks, J. 1974, "Capital Controversies: Ancient and Modern", *The American Economic Review*, vol. 64, no. 2, Papers and Proceedings of the Eighty-sixth Annual Meeting of the American Economic Association, pp. 307-316.
- Hicks, J. 1967, *Critical Essays in Monetary Theory*, Clarendon Press, Oxford.
- Hilferding, R. 1910, *Finance Capital: A Study of the Latest Phase of Capitalist Development*, Routledge & Kegan Paul, London, Boston, Melbourne and Henley.
- Ilyenkov, E.V. 2008 [1960], *The Dialectics of the Abstract and the Concrete in Marx's Capital*, Aakar Books, Delhi.
- Inklaar, R. & Wang, J.C. 2007, *Not Your Grandfather's Bank Any More? Consistent Measurement of Non-Traditional Bank Output*, Mimeo.

- Itoh, M. 1988, *The Basic Theory of Capitalism: The Forms and Substance of the Capitalist Economy*, Macmillan Press, Houndmills, London.
- Itoh, M. & Lapavistas, C. 1999, *Political Economy of Money and Finance*, Macmillan Press Ltd, Houndmills, London.
- Kalecki, M. 1990 [1943], "Political Aspects of Full Employment" in *Collected Works of Michal Kalecki*, ed. J. Osiantynski, Oxford University Press, Oxford.
- Kaufman, G.G. & Mote, L.R. 1994, "Is banking a declining industry? A historical perspective", *Economic Perspectives*, pp. 2-21.
- Kerr, G. 2011, *The Law of Opposites: Illusory Profits in the Financial Sector*, Adam Smith Institute, London.
- Keynes, J.M. 1973 [1936], *The General Theory of Employment, Interest, and Money. The Collected Writings of John Maynard Keynes. Vol. VII*, Macmillan Press for the Royal Economic Society, London.
- Keynes, J.M. 1971 [1930], *A Treatise on Money. The Pure Theory of Money. The Collected Writings of John Maynard Keynes. Vol. V*, Macmillan, Cambridge University Press, London, Basingstoke, New York.
- Keynes, J.M. 1939, "The Process of Capital Formation", *The Economic Journal*, vol. 49, no. 195, pp. 569-574.
- Keynes, J.M. 1938, "A Reply to Robertson", *The Economic Journal*, vol. 48, no. 190, pp. 318-322.
- Keynes, J.M. 1937a, "The "Ex-Ante" Theory of the Rate of Interest", *The Economic Journal*, vol. 47, no. 188, pp. 663-669.
- Keynes, J.M. 1937b, "Alternative Theories of the Rate of Interest", *The Economic Journal*, vol. 47, no. 186, pp. 241-252.
- Keynes, J.M. 1964 [1936], *The General Theory of Employment, Interest, and Money*, Harcourt.
- Knight, M.D. 2008, "Some reflections on the future of the originate-to-distribute model in the context of the current financial turmoil", *Speech at the Euro 50 Group Roundtable on "The future of the originate and distribute model"*.
- Kopke, W.K. 1982, "Forecasting Investment Spending: The Performance of Statistical Models", *New England Economic Review*, vol. November-December, pp. 13-29.

- Kosnik, R. & Shapiro, L. 1997, "Agency Conflicts between Investment Banks and Corporate Clients in Merger and Acquisition Transactions: Causes and Remedies", *The Academy of Management Executive*, vol. 11, no. 1, pp. 7-20.
- Kotz, D.M. 2011, "Financialization and Neoliberalism" in *Relations of Global Power: Neoliberal Order and Disorder*, eds. G. Teeple & S. McBride, University of Toronto Press, Toronto, pp. 1-18.
- Kotz, D.M. 2009, "The Financial and Economic Crisis of 2008: A Systemic Crisis of Neoliberal Capitalism", *Review of Radical Political Economics*, vol. 41, no. 3, pp. 305-317.
- Kotz, D.M. 1978, *Bank Control of Large Corporations in the United States*, University of California Press, Berkeley, Los Angeles, London.
- Kregel, J. 2010, "No Going Back: Why We Cannot Restore Glass-Steagall's Segregation of Banking and Finance", *Public Policy Brief 107*.
- Krippner, G.R. 2005, "The Financialization of the American Economy", *Socioeconomic Review*, vol. 3, pp. 173-208.
- Lamoreaux, N.R. 1985, *The Great Merger Movement in American Business, 1895-1904*, Cambridge University Press, Cambridge, New York, Melbourne.
- Lapavitsas, C. 2007, "Information and Trust as Social Aspects of Credit", *Economy and Society*, vol. 36, no. 3, pp. 416-436.
- Lapavitsas, C. 2004, "Hilferding's Theory of Banking in the Light of Stuart and Smith" in *Neoliberalism in Crisis, Accumulation, and Rosa Luxemburg's Legacy (Research in Political Economy)*, ed. P. Zarembka, Emerald Group Publishing Limited, , pp. 161-180.
- Lapavitsas, C. & Levina, I. 2011, "Financial Profit: Profit from Production and Profit upon Alienation", *Research on Money and Finance Discussion Paper no 24*.
- Lapavitsas, C. 2009, "Financialized Capitalism: Crisis and Financial Expropriation", *Historical Materialism*, vol. 17, no. 2, pp. 114-148.
- Lapavitsas, C. 2003, *Social Foundations of Markets, Money and Credit*, Routledge, London, New York.
- Lapavitsas, C. 2000, "On Marx's Analysis of Money Hoarding in the Turnover of Capital", *Review of Political Economy*, vol. 12, no. 2, pp. 219-235.
- Lazonick, W. 2007, "The US stock market and the governance of innovative enterprise", *Industrial and Corporate Change*, vol. 16, no. 6, pp. 983-1035.

- Lazonick, W. 2004, "Corporate Restructuring" in *The Oxford Handbook of Work and Organization*, eds. S. Ackroyd, R. Batt, P. Thompson & P. Tolbert, Oxford University Press.
- Lazonick, W. 1992, "Controlling the Market for Corporate Control: The Historical Significance of Managerial Capitalism", *Industrial and Corporate Change*, vol. 1, no. 3, pp. 445-488.
- Leland, H.E. & Pyle, D.H. 1977, "Informational asymmetries, financial structure, and financial intermediation", *The Journal of finance*, vol. 32, no. 2, pp. 371-387.
- Lunghini, G. & Bianchi, C. 2004, "The Monetary Circuit and Income Distribution: Bankers as Landlords?" in *Money, Credit, and the Role of the State: Essays in Honour of Augusto Graziani*, eds. R. Arena & N. Salvadori, Ashgate, Aldershot, Burlington, pp. 152-174.
- Martynova, M. & Renneboog, L. 2008, "A Century of Corporate Takeovers: What Have We Learned and Where Do We Stand?", *Journal of Banking & Finance*, vol. 32, no. 10, pp. 2148-2177.
- Marx, K. 1991 [1861-63], *Economic Manuscript of 1861-63. Marx-Engels Collected Works. Vol. 33*, International Publishers, New York.
- Marx, K. 1972 [1905-10], *Theories of Surplus Value, Part III*, Lawrence & Wishart, London.
- Marx, K. 1969 [1905-10], *Theories of Surplus Value, Part I*, Lawrence & Wishart, London.
- Marx, K. 2003 [1867], *Capital, Volume I*, International Publishers, New York.
- Marx, K. 1977 [1894], *Capital, Volume III*, International Publishers, New York.
- Marx, K. 1967 [1885], *Capital, Volume II*, International Publishers, New York.
- Mattey, J. & Wallace, N. 1999, *Housing Price Cycles and Prepayment Rates of U.S. Mortgage Pools*, Federal Reserve Bank of San Francisco.
- Mehrling, P. 2011, *The New Lombard Street: How the Fed Became the Dealer of Last Resort*, Princeton University Press, Princeton, Oxford.
- Mehrling, P. 2006, "The Problem of Time in the DSGE Model and the Post Walrasian Alternative" in *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, ed. D. Colander, Cambridge University Press, New York, pp. 70-79.

- Mehrling, P. 2001, "Love and Death: The Wealth of Irving Fisher" in *A Research Annual (Research in the History of Economic Thought and Methodology, Volume 19)* Emerald Group Publishing Limited, pp. 47-61.
- Mehrling, P. 1996, "The Relevance to Modern Economics of the Banking School View" in *Money in Motion: The Circulation and Post Keynesian Approaches*, eds. E. Nell & G. Deleplace, Macmillan, London, pp. 331-340.
- Mester, L.J. 1992, "Traditional and nontraditional banking: An information-theoretic approach", *Journal of Banking & Finance*, vol. 16, no. 3, pp. 545-566.
- Miller, G.P. 1998, "On the Obsolescence of Commercial Banking", *Journal of Institutional and Theoretical Economics*, vol. 154, no. 1, pp. 61-73.
- Mink, R. 2008, *An Enhanced Methodology of Compiling Financial Intermediation Services Indirectly Measured (FISIM)*, paper presented at OECD Working Party on National Accounts, Paris, 14-16 October 2008, available at [http://www.oilis.oecd.org/olis/2008doc.nsf/LinkTo/NT000059AE/\\$FILE/JT03251258.PDF](http://www.oilis.oecd.org/olis/2008doc.nsf/LinkTo/NT000059AE/$FILE/JT03251258.PDF).
- Minsky, H.P. 2008 [1987], "Securitization", *Policy Note No 2, Levy Economics Institute of Bard College*.
- Minsky, H.P. 2008 [1986], *Stabilizing an Unstable Economy*, McGraw-Hill, New York.
- Minsky, H.P. 1994, "Financial Instability and the Decline (?) of Banking: Public Policy Implications", *Levy Economics Institute Working Paper No. 127*.
- Minsky, H.P. 1993, "Finance and Stability: The Limits of Capitalism", *Working Paper No 93 Levy Economics Institute of Bard College*.
- Minsky, H.P. 1990, "Schumpeter: Finance and Evolution" in *Evolving Technology and Market Structure: Studies in Schumpeterian Economics*, eds. A. Heertje & M. Perlman, University of Michigan Press, Ann Arbor, pp. 51-74.
- Minsky, H.P. 1986, "The Evolution of Financial Institutions and the Performance of the Economy", *Journal of Economic Issues*, vol. 20, no. 2, pp. 345-353.
- Minsky, H.P. 1975, *John Maynard Keynes*, Columbia University Press, New York.
- Minsky, H.P. 1981, "James Tobin's Asset Accumulation and Economic Activity: A Review Article", *Eastern Economic Journal*, vol. 7, no. 3/4, pp. pp. 199-209.
- Mishkin, F.S. 2008, "On "Leveraged Losses: Lessons from the Mortgage Meltdown"", *Speech at the U.S. Monetary Policy Forum, New York, New York, February 29, 2008*.

- Mishkin, F.S. 2001, *The economics of money, banking, and financial markets*, 6th edn, Addison Wesley, Boston.
- Moore, B.J. 1988, *Horizontalists and Verticalists. The Macroeconomics of Credit Money*, Cambridge University Press, Cambridge, New York.
- Moulton, H.G. 1918, "Commercial Banking and Capital Formation: III", *Journal of Political Economy*, vol. 26, no. 7, pp. 705-731.
- Myers, S.C. & Rajan, R.G. 1998, "The Paradox of Liquidity", *The Quarterly Journal of Economics*, vol. 113, no. 3, pp. 733-771.
- Navin, T.R. & Sears, M.V. 1955, "The Rise of a Market for Industrial Securities, 1887-1902", *The Business History Review*, vol. 29, no. 2, pp. 105-138.
- Nell, E.J. 2001, "Notes on Hicks on Money and Monetary Theory" in *Credit, Interest Rates and the Open Economy: Essays on Horizontalism*, eds. L. Rochon & M. Vernengo, Edward Elgar, Cheltenham, Northampton, pp. 177-192.
- Orhangazi, O. 2007, *Financialization and the US Economy*, Edward Elgar, Northampton.
- Parguez, A. 2006, "The Absolute Paradox of Economic Policy in Contemporary Capitalism" in *Post-Keynesian Principles of Economic Policy*, eds. C. Gnos & L. Rochon, Edward Elgar, Cheltenham, Northampton, pp. 135-163.
- Parguez, A. 2004, "Moneda y Credito en el Capitalismo Contemporáneo (Money and Credit in Contemporary Capitalism)" in *Economía Financiera Contemporánea*, eds. E. Correa & A. Girón, Porrula Miguel Angel, pp. 15-54.
- Parguez, A. 2002, "Victoria Chick and The Theory of The Monetary Circuit: An enlightening debate" in *Money, Macroeconomics and Keynes: Essays in Honour of Victoria Chick*, eds. P. Arestis, M. Desai & S. Dow, Routledge, London, New York, pp. 45-55.
- Parguez, A. 2001, "Money without Scarcity: From the Horizontalist Revolution to the Theory of the Monetary Circuit" in *Credit, Interest Rates, and the Open Economy: Essays on Horizontalism*, eds. L. Rochon & M. Vernengo, Edward Elgar, Cheltenham, Northampton, pp. 69-103.
- Parguez, A. 1987, "Introduction a l'economie de rentiers", *Economies et Societes*, vol. 21, no. 9, pp. 103-119.
- Parguez, A. & Seccareccia, M. 2000, "The Credit Theory of Money: The Monetary Circuit Approach" in *What is Money?*, ed. J. Smithin, Routledge, London, New York, pp. 101-123.

- Phillips, C.A. 1920, *Bank credit: A study of the Principles and Factors Underlying Advances Made by Banks to Borrowers*, Macmillan, New York.
- Pollin, R. 2007, "Resurrection of the Rentier", *New Left Review*, vol. 46, pp. 140-153.
- Pollin, R. 1996, "Contemporary Economic Stagnation in World Historical Perspective", *New Left Review*, vol. I, no. 219, pp. 109-118.
- Pozsar, Z., Adrian, T., Ashcraft, A. & Boesky, H. 2010, "Shadow Banking", *Federal Reserve Bank of New York Staff Report no. 458*, .
- Radecki, L.J. 1999, "Banks' Payments-Driven Revenues", *FRBNY Economic Policy Review*, pp. 53-70.
- Rajan, R.G. 1996, "Why Banks Have A Future: Toward A New Theory Of Commercial Banking", *Journal of Applied Corporate Finance*, vol. 9, no. 2, pp. 114-128.
- Reid, J. 2008, *A Trillion-Dollar Mean Reversion?*, Deutsche Bank.
- Ricardo, D. 1821, *On the Principles of Political Economy, and Taxation*, John Murray, London.
- Robertson, D.H. 1937, "Alternative Theories of the Rate of Interest: Three Rejoinders", *The Economic Journal*, vol. 47, no. 187, pp. 428-436.
- Robinson, J. 1985, "The Theory of Normal Prices and Reconstruction of Economic Theory" in *Issues in Contemporary Macroeconomics and Distribution*, ed. G.R. Feiwel, State University of New York Press, New York.
- Rochon, L. & Rossi, S. 2010, "Has "It" Happened Again?", *International Journal of Political Economy*, vol. 39, no. 2, pp. 5-9.
- Rossi, S. 2003, "Money and Banking in a Monetary Theory of Production" in *Modern theories of money: the nature and role of money in capitalist economies*, eds. L. Rochon & S. Rossi, Edward Elgar, Cheltenham, Northampton, pp. 339-359.
- Rotta, T.N. & Teixeira, R.A. 2011, *A Marxian Theory of Financialization*, Paper presented at the Eastern Economic Association Annual Conference, New York City, February 2011.
- Samolyk, K. 2004, *The Future of Banking in America. The Evolving Role of Commercial Banks in U.S. Credit Markets*.
- Schmidt, R.H. 1999, "Disintermediation and the Role of Banks in Europe: An International Comparison", *Journal of financial intermediation*, vol. 8, no. 1-2, pp. 36.

- Scholtens, B. 2000, "A critique on the theory of financial intermediation", *Journal of banking finance*, vol. 24, no. 8, pp. 1243.
- Scholtens, B. & Wensveen, D.v. 2003, "The Theory of Financial Intermediation: An Essay On What It Does (Not) Explain", *SUERF – The European Money and Finance Forum*.
- Schumpeter, J.A. 1961 [1911/1934], *The Theory of Economic Development*, Oxford University Press, New York.
- Schumpeter, J.A. 1954, *History of Economic Analysis*, Routledge, London.
- Sekine, T.T. 1983, *The Dialectic of Capital: A Study of the Inner Logic of Capitalism*, York University.
- Shleifer, A. & Vishny, R.W. 2003, "Stock market driven acquisitions", *Journal of Financial Economics*, vol. 70, no. 3, pp. 295-311.
- Skidelsky, R. 1992, *John Maynard Keynes, Volume two. The Economist as Saviour. 1920-1937*, Macmillan, London.
- Smith, A. 1884, *An Inquiry into the Nature and Causes of the Wealth of Nations*, T. Nelson and Sons, London.
- Steuart, J. 1770a, *An Inquiry into the Principles of Political Economy, Vol. I*, James Williams and Richard Moncrieffe, Dublin.
- Steuart, J. 1770b, *An Inquiry into the Principles of Political Economy, Vol. II*, James Williams and Richard Moncrieffe, Dublin.
- Stiroh, K.J. 2004, "Diversification in Banking: Is Noninterest Income the Answer?", *Journal of Money, Credit and Banking*, vol. 36, no. 5, pp. 853-882.
- Stockhammer, E. 2010, "Financialization and the Global Economy", *PERI Working Paper No. 240*.
- Stockhammer, E. 2004, "Financialization and the Slowdown of Accumulation", *Cambridge Journal of Economics*, vol. 28, pp. 719-741.
- Teixeira, R.A. & Rotta, T.N. 2010, Modern Rent-Bearing Capital and Financialization: New Enclosures, Knowledge-Rent and Valueless Commodities, *Mimeo*.
- Thornton, H. 1802, *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain*, J. Hatchard, London.

- Tirole, J. 2011, "Illiquidity and All Its Friends", *Journal of Economic Literature*, vol. 49, no. 2, pp. 287-325.
- Tobin, J. 1969, "A General Equilibrium Approach To Monetary Theory", *Journal of Money, Credit and Banking*, vol. 1, no. 1, pp. 15-29.
- Tobin, J. 1963, "Commercial Banks as Creators of 'Money'", *Cowles Foundation Discussion Paper No 159*.
- Tobin, J. & Brainard, W. 1990, "On Crotty's Critique of q-Theory", *Journal of Post Keynesian Economics*, vol. 12, no. 4, pp. 543-549.
- Turner, A. 2010, "What do banks do? Why do credit booms and busts occur and what can public policy do about it?" in *The Future of Finance: The LSE Report*. London School of Economics and Political Science, pp. 5-86.
- van Treeck, T. 2009, "The Political Economy Debate on 'Financialization' – A Macroeconomic Perspective", *Review of International Political Economy*, vol. 16, no. 5, pp. 907-944.
- Veblen, T. 1904, *The Theory of Business Enterprise*, Charles Scribner's Sons, New York.
- Waldrop, R. 2002, *New Reporting Offers Insight into Bank Activities*, FDIC.
- Walter, T.S., Yawson, A. & Yeung, C.P.W. 2008, "The Role of Investment Banks in M&A Transactions: Fees and Services", *Pacific-Basin Finance Journal*, vol. 16, no. 4, pp. 341-369.
- Wang, J.C. 2003, *Loanable Funds, Risk, and Bank Service Output*, Federal Reserve Bank of Boston Working Paper Series No. 03 - 4, Boston.
- Wang, J.C., Basu, A. & Fernald J.G. 2004, *A General-Equilibrium Asset-Pricing Approach to the Measurement of Nominal and Real Bank Output*, Invited for conference volume on Price Index Concepts and Measurement, Conference on Research on Income and Wealth (CRIW), available at <http://www.bos.frb.org/economic/wp/wp2004/wp047.htm>.
- Wang, C. & Uryasev, S. 2006, *Efficient Execution in the Secondary Mortgage Market: A Stochastic Optimization Model Using CVaR Constraints*, Department of Industrial and Systems Engineering, University of Florida.
- Willis, H.P. 1921, *American Banking*, La Salle Extension University, Chicago.
- Wray, L.R. 2011, *Minsky's Money Manager Capitalism and the Global Financial Crisis*, Working Paper No. 661, Levy Economics Institute of Bard College.

Wray, L.R. 2009, "The Rise and Fall of Money Manager Capitalism: A Minskian Approach", *Cambridge Journal of Economics*, vol. 33, no. 4, pp. 807-828.

Xu, X.E. & Fung, H. 2005, "What Moves the Mortgage-Backed Securities Market?", *Real Estate Economics*, vol. 33, no. 2, pp. 397-426.