UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Faculty of Commerce, Law and Management

Candidate's Declaration

Name of Candidate:	NICHOLAS SPEARMAN
Person Number:	0710274 T
Degree:	CD000 - DOCTOR OF PHILOSOPHY
Supervisor / Co-Supervisor:	PROF GREG FARRELL
School:	SCHOOL OF ECONOMICS
Title of Thesis/Dissertation/ Research Report:	MONETARY INTERMEDIATION THEORY: THE ROLE OF BANKS AND DEBT IN THE MONETARY SYSTEM.

Candidate's Declaration:

ii. I hereby submit my PhD Thesis/Masters dissertation for examination (circle applicable).

iii. I confirm that my signed declaration in terms of Rule G9.7 is included in each copy of the thesis/ dissertation.

iv. I have checked all copies of my thesis/ dissertation and declare that no pages are missing or poorly reproduced.

I hereby submit a CD/USB containing a PDF version of my submission for examination and 1 bound copies.

Candidate's Signature: ______________ **Date:** 5 DECEMBER 2019

Monetary Intermediation Theory

On the role of banks and debt in the monetary system

NC Spearman

2020

"Perhaps no other family left such an imprint on an age as the Medici left on the Renaissance. ... They came to be known as bankers (*banchieri*) because, like the Jews of Venice, they did their business literally seated at benches [*banci*] behind tables in the street."

— Ferguson (2009, pp. 42-43)

"The popular notion is that the word Bank is derived from the Italian word *banco*, a bench or table, because the early Italian money dealers, or money changers, displayed their specie on a bench or table ... whence they were called *banchieri.* ... Nevertheless, there can be no possible doubt but that this derivation is a pure delusion. In the first place the money dealers, or money changers, as such, were never called *banchieri* in the middle ages ... and their places of business were called *casane*, and not *banchi.*"

— Macleod (1866, pp. 109-111)

“When money was made of gold or other metals, it was inconvenient for consumers and merchants to carry it around and weigh and assay its purity every time they made a transaction. So the practice developed of leaving gold in a goldsmith’s safe storage facilities and carrying in its place a receipt … . When people began trading goods and services for the goldsmiths’ receipts, rather than for the gold itself, the receipts became an early form of paper money … . [From this] the modern system of **fractional reserve banking** was born”

— Baumol & Blinder (2011, pp. 626-627); undergraduate textbook

“The usual textbook view is that banking developed from goldsmiths who issued receipts for gold left with them which later circulated from hand to hand, and that observation of this circulation ultimately induced goldsmiths to issue receipts without previous deposit. The story is well told, but inaccurate.”

— Kindleberger (1984, p. 35)

"[P]eople deposit money in banks, which in turn lend this money to people like you and me."

— Greenbaum, et al. (2016, p. 24)

"Banking is not money lending; to lend, a money lender must have money. … Banks are important exactly because they do not operate under the constraint of a money lender"

— Minsky (2008 [1986], p. 256; p. 278)

"[I]t is with peculiar diffidence and even apprehension that one ventures to open one's mouth on the subject of money."

— Hicks (1935, p. 1)

Thank you

The business of banks and their connection to the subject of money remains mired in centuries of controversy. With this thesis, I wade into the furore. To every-(long-suffering)-one whose support has made this thesis possible—family, friends, colleagues, and my ever-patient supervisor—for your willingness to endure stories about "what banks do" and for your subsequent jokes about cash-money, I am deeply grateful. The privilege of dedicating time and intellect to this fascinating subject has been intensely rewarding. To anyone with an interest in the institution of banking and its entanglement with the phenomenon of money and the effects on the economy more broadly, I sincerely hope you find value in the discussion on the following pages.

Abstract

Why are banks central to the monetary system? And how does their role in this system affect the real economy? In the wake of the 2007/8 financial crisis, these fundamental questions have re-surfaced as both controversial and crucial to the study of macroeconomics. Interest in this questioning has been buoyed by the wave of excitement that accompanied the "crypto-currency" bubble of 2017; and by the promises of emerging "fintech" innovations premised on the idea of revolutionising money and banking. In this thesis, I present a novel framework for thinking about the role of banks in the economy and make three contributions to the literature. First, I develop a *monetary* theory of banks, and show how this is useful for explaining why bank liabilities are central to the monetary system; for explaining bank balance sheet structure, products, and contracts; and for providing new and compelling answers to other enduring puzzles within bank theory. Second, I test empirical evidence from South African bank balance sheet data and find support for the monetary intermediation theory of banks. I discuss the policy implications of these findings, as well as the broader implications of framing banks as monetary intermediaries more generally. Third, I extend the dynamic stochastic general equilibrium framework to incorporate the essential features of monetary intermediation and analyse the affects. The model indicates that monetary intermediation matters for real economic activity. Credit shocks are found to generate real instability, and endogenous inflation targeting monetary policy is shown to be an effective tool for mitigating the real effects of these shocks. The results contribute to contemporary debate concerning the role of both banks and monetary policy in the economy.

Table of Contents

List of figures

List of tables

List of abbreviations

AD	–	aggregate demand
AS	–	aggregate supply
CBDC	–	central bank digital currency
CIA	–	cash-in-advance
CRRA	–	constant relative risk aversion
CPI	–	consumer price index
DSGE	–	dynamic stochastic general equilibrium
FEVD	–	forecast-error variance decomposition
FOC	–	first-order condition
G&S	–	goods and services
GDP	–	gross domestic product
IC	–	information criteria
IES	–	intertemporal elasticity of substitution
IRF	–	impulse-response function
MMF	–	money market fund
NK	–	New Keynesian
OLG	–	overlapping generations
RBC	–	real business cycle
SABOR	–	South African bank overnight rate
SARB	–	South African Reserve Bank
StatsSA	–	Statistics South Africa
SVAR	–	structural vector autoregression
TFP	–	total factor productivity
VAR	–	vector autoregression

Introduction

> "I should like then in all humility to call attention to the fact that the banks' prime duty is not to earn a great deal of money but to provide the public with a medium of exchange ..."
>
> — Wicksell (1936 [1898], p. 190)

I. Motivation

Money has a unique and interesting property. People accept it as payment for the things they sell (e.g., labour), and they use it to pay for the things they buy in return (e.g., food, shelter, and entertainment). In other words, buyers and sellers use money as a generally accepted *medium-of-exchange*; meaning it is a *medium* by which one lot of goods and services (G&S) can be indirectly *exchanged* for another lot of G&S. All money has this property; or at least colloquially, anything that has this property is considered to *be* money—that is, this property is regarded as a *distinguishing* feature of money.[1]

[1] Money is commonly defined according to its functional roles. These are traditionally identified as: a measure-of-value (also referred to as a unit-of-account), a generally accepted medium-of-exchange, a store-of-value, and a standard of deferred payment. Schumpeter (2006 [1954], p. 1053) refers to these as the "old four functions", having been discussed by the likes of Plato and Aristotle, and by many others since. But as Angell (1930, p. 69) writes: "By 'money' economists usually mean anything that is (1) passed from hand to hand in payment for commodities and service, and (2) regularly taken with the intention of offering it in payment to others, and (3) customarily received without assay or special test of quality or quantity, and (4) received without reference to our reliance upon the personal credit of the one who offers it". The description emphasises the function of money as a medium-of exchange. Likewise, von Mises (1953 [1912], p. 34) writes: "The simple statement, [is] that money is a commodity whose economic function is to facilitate the interchange of goods and services". Although this view is not universal, Borio (2019) describes the medium-of-exchange function of money as its "distinguishing feature".

Numerous artefacts from across diverse cultures and environments (including shells, carved stones, baked clay tokens, and ingots of shaped metal) have historically functioned as generally accepted media-of-exchange.[2] But money supply today consists chiefly of two types of *financial instruments* that serve this purpose. The first are the notes and coins issued by state entities—these are referred to within economic theory as fiat money, outside money, exogenous money, currency, or cash. The second are the demand-deposit balance sheet liabilities of private firms called banks—these liabilities are referred to as inside money, endogenous money, credit-money, or bank-money. This makes banks special; the liabilities of other types of firms do not function as money. Why is this? What is it about banks that make them so special that they are central to the monetary system? And how does this affect the economy? These questions remain remarkably unresolved despite a history of banking that stretches back thousands of years. It is this broad line of inquiry that motivates this thesis.

II. Context

Following the financial crisis of 2007/8, there has been growing interest in, and calls for, more deeply considering the role of banks and debt in the monetary system and the economy more broadly.[3] Contemporary bank theory falls largely within two distinct areas of economic enquiry: financial intermediation theory on the one hand and monetary economics on the other.

[2] See Angell (1930) and Davies (2010) for discussions on historical artefacts used as media-of-exchange.
[3] See for example Gorton (2014), Borio (2014; 2019), and Decker & Goodhart (2018).

From the perspective of financial intermediation theory, banks mitigate frictions that impair the flow of funds from lenders to borrowers (Gurley & Shaw, 1960; Tobin, 1963; Fama, 1980; Greenbaum, et al., 2016). The primary role of banks in this regard is the provision of portfolio management services.[4] Financial intermediation theory is therefore concerned with the asset function of bank liabilities as a store-of-value, and says little about the unique transactions function of bank liabilities as a medium-of-exchange.[5] This is perhaps surprising given that the use of bank liabilities as a medium-of-exchange is recognised as a unique and defining characteristic of banks within financial intermediation literature.[6]

By contrast, monetary economics focuses on the transaction services of commodities and financial instruments that serve as media-of-exchange. Banks are therefore of interest because their liabilities function as generally accepted media-of-exchange that provide transaction services. Some recent studies focused on bank transaction services argue that banks play no intermediation role at all.[7]

But dissonance between money and finance theories is a long-standing phenomenon. This is lamented by Lucas (1984).[8] Kiyotaki & Moore (2002) acknowledge

---

[4] "[Banks] provide transactions services … , and they provide portfolio management services … ; the concern with banks in macroeconomics centers on their role as portfolio managers" (Fama, 1980, p. 44).

[5] Tobin (1963, p. 16) writes: "The distinction between commercial banks and other financial intermediaries has been too sharply drawn. The differences are of degree, not of kind … . In particular, the differences which do exist have little intrinsically to do with the monetary nature of bank liabilities".

[6] Gorton & Winton (2003, pp. 434-5) write: "Demand deposits are securities with special features … securities that have the important property of being a medium of exchange".

[7] See as examples de Soto (2009), Benes & Kumhof (2012), McLeay, et al. (2014a; 2014b), Werner (2014a), and Jakab & Kumhof (2015). Jakab & Kumhof (2015, p. 3) write: "whenever a bank makes a new loan to a non-bank customer X, it creates a new loan entry in the name of customer X on the asset side of its balance sheet, and it simultaneously creates a new and equal-sized deposit entry, also in the name of customer X, on the liability side of its balance sheet … And because both entries are in the name of customer X, there is no intermediation whatsoever".

[8] "[T]he objective of constructing a unified theory of money and finance is an old one, one that has challenged theorists at least since J.R. Hicks" (Lucas, 1984, p. 9).

that reconciling monetary and finance theories is difficult.[9] Brunnermeier & Sannikov (2016) acknowledge that monetary theory needs a proper place for financial intermediaries, and this is the focus of their "I-theory" of money. But the I-theory focuses on bank deposits as a riskless store-of-value, and so does little to reconcile financial intermediation theory with the *transactions* focus of monetary economics. There is, however, a growing body of literature aimed at integrating the role of banks as risk- and cost-mitigating *intermediaries* with the use of their liabilities as a broadly circulating and generally accepted *medium-of-exchange*. These studies focus on how G&S are exchanged backwards and forwards through *time*.[10]

For example, Bullard & Smith (2003b) present a model in which agents transact by pledging G&S to be delivered in the *future*, in exchange for G&S received *now*. These G&S are traded through time using debt-credit contracts.[11] The debt-credit contracts are the media by which the G&S are traded. Bullard & Smith show that if these contracts are subject to transaction costs, and if contracting with certain intermediary agents reduces these costs, then these intermediary contracts will displace costlier bilateral contracts to become ubiquitous media-of-exchange, and thereby improve the efficiency of trade.

Similarly, Kiyotaki & Moore (2002) present a model in which agents trade G&S across time using debt-credit contracts. But not all agents are able to *credibly* commit to

[9] "To justify why agents cannot simply trade their commodities through a market, there must be some physical trading friction. Search or matching frictions are often invoked, but unfortunately noncompetitive models of this kind, though ingenious and elegant, are difficult to incorporate into a standard macroeconomic framework" (Kiyotaki & Moore, 2002, p. 62).

[10] In other words, by contrast to the barter theories of money that focus on exchanging different types of G&S with the *same* time characteristics, these studies focus specifically on *differences* in the time characteristics of G&S exchanged.

[11] I use the term "debt-credit" for what in economic literature is referred to interchangeably as either "debt" or "credit" and which therefore has the potential for confusion. As Hawtrey (1919) states: "[t]hat which to the debtor is a debt, is to the creditor a credit", and hence any debt-credit contract has simultaneously both debt and credit aspects. If I do not intend to emphasise either the debt or the credit aspects of these contracts, I use the term debt-credit instead.

debt obligations. This prevents certain transactions and hinders efficient trade. Kiyotaki & Moore show that if debt-credit contracts are *transferable*, then contracts issued by *credibly* trustworthy agents can mitigate these inefficiencies by enabling less credible agents to acquire them and transfer them in subsequent transactions.

Williamson & Wright (2011) present a model in which agents trade G&S using debt-credit contracts, and, like Kiyotaki & Moore, information frictions mean that in some instances agents are unable to credibly commit to debt obligations. The model includes cash, and cash enables buyers to transact even if they cannot credibly commit. However, holding cash has an opportunity cost: it is unproductive, whereas G&S that can be pledged can be productively invested until transferred. In the model, agents have access to a third-party technology (a bank) that enables them to collectively both store their cash and invest their G&S productively in a *combined pool* of assets. Agents are issued with "deposit" receipts. These receipts are useful because they can be pledged when agents are able to credibly commit their invested G&S; and when unable to commit credibly they can be liquidated into cash. This arrangement improves efficiency by providing agents with Diamond-Dybvig type *risk sharing* which reduces the opportunity costs of agents holding cash individually (Diamond & Dybvig, 1983).

Contemporary literature connecting the use of bank liabilities as a medium-of-exchange with the role of banks as friction-mitigating intermediaries represents renewed interest in a topic central to earlier bank theory as reflected in the literature of the 19th and early 20th centuries. This earlier literature is specific in its focus on the role of banks as special intermediaries whose liabilities circulate as a generally accepted medium-of-exchange. This literature highlights that bank balance sheets are designed to mitigate the risks and costs of exchanging G&S on the basis of debt-credit contracts.[12] Bank liabilities are particularly important within this literature because they provide a conduit for promoting real economic activity. Prominent contributors include Thornton, (1965 [1802]), Tooke (1844), Macleod (1866; 1889), Bagehot (1873), Wicksell (1936 [1898]), von Mises (1953 [1912]), Keynes (1930), and Schumpeter (1983 [1934]).

Sissoko (2016) develops a contemporary model of banking based on this early theory. Like Kiyotaki & Moore (2002) and Williamson & Wright (2011), agents in Sissoko's model exchange G&S across time, and commitment problems hamper the use of bilateral debt-credit contracts as media-of-exchange. Cash is available, and certain agents are able to provide both credible commitment and are able to monitor debt obligations. These agents (banks) purchase bilateral debt-credit contracts at a discount and provide their own trusted and readily transferable liabilities as payment. These liabilities circulate as a widely accepted medium-of-exchange and facilitate further trade. The unique feature of the model is that cash is not only costly to hold (because it pays no interest), but it is also scarce; and transferable bank liabilities provide a cost minimising payment *substitute* for this scarce cash. Therefore, unlike Williamson & Wright or Kiyotaki & Moore, Sissoko's model captures the role of banks as friction-mitigating intermediaries whose liabilities function as a *direct cash substitute*. This feature makes banks central to Sissoko's monetary system.

[12] This largely corresponds to what Decker & Goodhart (2018) refer to as the theory of "credit mechanics".

III. Contribution

Following the early literature of Bagehot (1873) and Wicksell (1936 [1898]) and more recently Sissoko (2016), this thesis is a contribution towards understanding why banks are central to the monetary system. I present a novel framework for thinking about the role of banks in the economy and make three contributions to the literature. First, I develop a theory of banks as monetary intermediaries and show how this is useful for explaining why bank liabilities are central to the monetary system; for explaining bank balance sheet structure, products, and contracts; and for providing new and compelling answers to other enduring puzzles within financial intermediation literature. Second, I test empirical evidence from South African bank balance sheet data and find support for the monetary intermediation theory of banks. I discuss the policy implications of these findings, as well as the broader implications of framing banks as monetary intermediaries more generally. Third, I extend the dynamic stochastic general equilibrium framework to incorporate the essential features of monetary intermediation and analyse the affects. The results contribute to contemporary debate concerning the role of banks and monetary policy in the economy.

Decker & Goodhart (2018) stress the importance of exploring alternative frameworks for understanding bank balance sheets to better understand money supply dynamics. With this thesis, I add to this discussion and also contribute to literature focused on the role of payments technology in macroeconomics.[13] There is also general interest in rethinking both monetary policy (Blanchard, et al., 2010) and financial stability (Aikman, et al., 2018). This thesis contributes to these dialogues by adding to

[13] See for example Kahn & Roberds (2009).

our understanding of the role of banks in the monetary system. This is pertinent for at least two reasons. The first is the increased emphasis on incorporating banking and other monetary sector mechanics into macroeconomic models following the crisis of 2007/8. The role of the monetary and banking systems in facilitating transactions within these models has by and large been taken for granted (Borio, 2019). I explore why banks play this role and what this means for monetary policy and economic activity in general. The second reason is the current plethora of emerging "fintech" innovations aimed at disrupting the operations and institutional structure of traditional financial systems. Clearly understanding a bank's monetary role is important in this context for understanding how banks may adapt to these innovations, how these emerging technologies may impact on policy transmission channels, and how these changes may be expected to shape future prudential regulations (Bech & Garratt, 2017; Engert & Fung, 2017; Julin, 2017; Löber & Houben, 2018).

Chapter 1

Bank Theory

"[A banker is like a friend who] loaned you an umbrella when the sun was shining and demanded its return the moment it started to rain"

— Mark Twain, cited in Cerf (1944, p. 241)

1.1. Introduction

Bank theory falls across three distinct bodies of economic theory: financial intermediation theory; monetary economics; and the theory of banks described in 19th and early 20th century banking literature. This first chapter provides a broad overview of these bodies of theory as they relate to this thesis. This serves to: (i) *situate* the theory of monetary intermediation within the broader body of bank theory literature; and (ii) *highlight questions, puzzles, and lacunae* within this literature that are addressed within monetary intermediation theory. This chapter is therefore structured as follows. Section 1.2 centres on the role of banks in financial intermediation theory. Section 1.3 reviews the role of intermediaries in monetary economics. Section 1.4 steps back in time to consider the role of banks described in 19th and early 20th century banking literature. Finally, Section 1.5 concludes by summarising how this thesis contributes to this bank theory literature.

1.2. Financial intermediation theory

There are three prominent reviews of the financial intermediation literature spanning successive decades: Santomero (1984), Bhattacharya & Thakor (1993), and Gorton & Winton (2003). Each reflects advances in theory and changes in emphasis that occurred during the intervening decades.[1] These contributions each summarise the major themes explored in the literature, as well as list unresolved issues for directing potential future

[1] Matthews & Thompson (2005), Mishkin & Eakins (2006), Freixas & Rochet (2008), and Greenbaum, et al. (2016) provide more recent and formal textbook treatments of this theory.

research. In addition, Merton (1993; 1995) and Merton & Bodie (1995; 2005) provide an alternative "functional perspective" viewpoint that differs in focus from the highlighted review studies.

In this section, I briefly review core themes highlighted by these studies that are pertinent to the theory developed in this thesis. Section 1.2.1 discusses market imperfections as the *raison d'être* of bank intermediation. This is followed by discussions on banks as liquidity providers (Section 1.2.2); the industrial organisation and behaviour of banks (Section 1.2.3); the functional perspective approach to financial intermediation theory (Section 1.2.4); and finally, Section 1.2.5 deals with the interconnected issues of bank regulation, financial stability, and bank panics.

1.2.1. Incomplete and imperfect markets as the raison d'être *of banks*

Literature on the role of incomplete and imperfect markets in financial intermediation theory can be divided into three approaches: (i) literature focusing on the asset side of a bank's balance sheet; (ii) literature focusing on the liabilities side of the balance sheet; and (iii) literature focusing on the interconnectedness of both bank assets and bank liabilities (Santomero, 1984).

Literature concerned with the asset side of a bank's balance sheet explores the role of financial intermediaries as asset managers. The underlying question is why do banks invest in assets on behalf of investors? This literature focuses on how deviations from the perfect-markets paradigm create space for financial intermediaries to provide *asset management services at lower costs than direct market participation*. Financial intermediation provides qualitative asset transformation through asset diversification and evaluation. Financially intermediated diversification is able to achieve a better risk-return profile for stakeholders and enables the transformation of large denomination financial assets into smaller denomination units at lower cost. Financially intermediated asset evaluation provides an asset quality filter that provides a signal of asset quality within an environment of imperfect information.

Literature concerned with the liabilities side of a bank's balance sheet explores the use of bank liabilities as a store-of-value asset class held by investors. The underlying question is why do investors hold bank deposits as an asset class? The literature focuses on how demand for deposits is determined as a function of liquidity needs, transaction costs, uncertainty with respect to underlying bank asset quality, and rates of return offered by banks. These determinants are evaluated against direct but costly and risky investment via market participation. However, a general lacuna within this literature concerns the answer to the following question: *why do the liabilities of certain institutions (banks) provide transaction services while other similar financial intermediaries do not*? This question is tackled in Chapter 2.

Literature concerned with the dual nature of bank balance sheets focuses on understanding why banks transform deposits with certain characteristics on the liabilities side of their balance sheet into loans with different characteristics on the asset side of their balance sheets. This literature generally argues that banks provide maturity and asset transformation services due to prohibitive market transaction costs and deviations from perfect-market assumptions which banks are able to exploit. Banks hold portfolios of loans and matching deposits because they can engage in risky arbitrage across markets. In this regard, the covariance of interest rates across the balance sheet is highlighted as a factor motivating the use of portfolio choice as an analytical tool for modelling financial intermediaries.

The common theme across all three approaches is that financial intermediation mitigates transaction costs that impede the transfer of funds from savers to investors. Bhattacharya & Thakor (1993) describe information asymmetries as the most basic transaction cost and the amelioration of "informational pathologies" as the fundamental reason for financial intermediation. This emphasis reflects the growth of interest in the role of information in economics following the seminal contributions of Akerlof (1970), Spence (1973), and Rothschild & Stiglitz (1976). Gorton & Winton (2003) highlight the importance of information asymmetry leading to issues such as delegated monitoring problems. Gorton & Winton describe how, if monitoring is sufficiently costly, it will be efficient to delegate the task to a specialised agent. Moreover, if information about investment opportunities is costly, then some economic agents may find it worthwhile to focus on producing this information. However, producing this information introduces both "appropriability" and "reliability" problems meaning that agents may not have the incentive to produce the information reliably. These problems introduce a need to "monitor the monitor"; and Gorton & Winton highlight how diversification, economies of scale, and the structure of bank contracts is able to reduce the costs of these problems.

The primary motivation for banks in financial intermediation theory is the mitigation of the risks and costs associated with lending funds to borrowers in environments with imperfect and incomplete markets. In this thesis, the theory presented is complementary, but the *emphasis differs in two important ways*. First, the focus on banks as monetary intermediaries concerns banks ameliorating various risks and costs associated with *exchanging G&S,* as opposed to *lending idle funds*. Second, the focus is not primarily on *information asymmetries*, and is framed instead in terms of more general Knightian *uncertainty*; meaning environments that are not limited to certainty about the possible probability distributions of future events (Knight, 1964 [1921]).

1.2.2. Risk sharing, liquidity provision, and maturity transformation

The role of banks in providing risk sharing and liquidity services through demand deposits features prominently in the literature review study of Bhattacharya & Thakor (1993)—presumably as a response to the savings and loan crisis in the US in the late 1980s. However, Bhattacharya & Thakor highlight that while demand deposits provide liquidity through risk sharing that counters random shocks to the timing of consumption, this leaves banks vulnerable to runs. Two studies that focus on this point and that are of particular relevance to this thesis are Diamond & Dybvig (1983) and Jacklin (1987).

In the classic Diamond-Dybvig model, consumers face an uninsurable consumption shock that results in early liquidation of investments and loss of potential investment returns. A bank provides consumption smoothing insurance against this risk through a demand-deposit contract, which allows early liquidation of deposits and at the same time attempts to provide an insured return. But the structure of this contract make banks vulnerable to runs. An early critique of the Diamond-Dybvig model by Jacklin (1987) demonstrates that dividend-paying equity shares provide the same risk-sharing allocation without the possibility of a bank run. Jacklin demonstrates that demand deposits enable greater risk sharing than trading in dividend-paying equity shares but only if certain restrictions are imposed on the trading of demand deposits.[2] If demand deposits can be freely transferred, dividend-paying shares provide the same risk-sharing allocation as demand deposits without the added possibility of a bank run. Jacklin therefore concludes that if financial markets are complete and frictionless, a demand-deposit intermediary can only exist to provide optimal risk sharing if individuals are limited by restrictions on the trading of their deposits. Haubrich & King (1990) make a similar point. Yet a defining feature of demand deposits is that they are readily and easily transferred. The result is therefore puzzling.

An important feature of the Diamond-Dybvig model is that deposit contracts do not function as money in the model even though consumers generally transfer bank liabilities directly as money without any need for liquidation (Gorton & Winton, 2003). This feature of bank deposits remains largely underexplored in the financial intermediation literature and the model in Chapter 2 focuses on precisely this point. The model is structured around the role of demand deposits as a directly and readily transferable medium-of-exchange. I show that this has important implications, because

[2] Jacklin (1987) shows that demand deposits allow greater risk sharing because they are a mechanism that promotes self-selection, but for self-selection to take place, allocations must be both individual incentive compatible and coalition incentive compatible. Without trading restrictions, trading opportunities devised by coalitions of agents can eliminate potential risk-sharing benefits. This is because alternative allocations available as a result of trading may not be *coalition* incentive compatible. In this case, individuals do not self-select and the mechanism breaks down.

in contrast to Jacklin (1987), demand deposits in the model provide optimal risk sharing *without trading restrictions*.

Bank deposits provide liquidity as a form of insurance that has desirable state contingent payoffs in the face of market imperfections (Gorton & Winton, 2003). The literature explores numerous models of this role.[3] Gorton & Winton distinguish three broad types of liquidity provided. The first is to provide instruments that are information insensitive in the face of asymmetric private information. The second is to provide state contingent liquidity in the face of moral hazard issues. The third is to provide agents who are spatially or temporally separated with the ability to contract and trade with one another. This third notion of liquidity features prominently in monetary theory literature (see Section 1.3), and it is central to this thesis. This thesis focuses on the role of banks in mitigating the cost of commitment problems in the face of uncertainty when agents trade with one another across time.

1.2.3. The industrial organisation of banks

Freixas & Rochet (2008) highlight the inability of the Arrow-Debreu (perfect and complete markets) paradigm to account for observed complexities of the banking industry. They identify two broad approaches to dealing with this. The first is the incomplete-markets paradigm (corresponding broadly to the discussion in Sections 1.2.1 and 1.2.2 above); and the second constitutes an industrial organisation approach. The industrial organisation approach overlaps in many instances with incomplete-market

[3] See as recent examples studies by He, et al. (2005), Donaldson, et al. (2016), Brunnermeier & Sannikov (2016), and Dang, et al. (2017).

explanations, but the focus is specifically on the internal organisation of the banking firms themselves and the services they offer. Banks are modelled as independent entities that optimally react to their environment, and the focus is on how their specific organisational characteristics influence equilibrium characteristics.

Santomero (1984) discusses the banking firm's optimisation problem as a function of stakeholder objectives for the firm. These stakeholders are typically regarded as equity investors or bank management, and their objectives influence the degree of risk aversion underlying the bank's objective function. This determines the bank's decisions concerning portfolio choices. Bhattacharya & Thakor (1993) discuss how factors activity the ownership structures of financial intermediaries; economies of scale and scope in financial intermediation activities; and the role of loan commitments as a contractual mechanism influence banking activities. Gorton & Winton (2003) provide extensive discussion on how the nature of a bank's relationships with its stakeholders shapes its industrial organisation, including delegated monitoring, information asymmetries, and bank contracts (especially in a dynamic setting). They conclude that the industrial organisation of the banking sector is a critical factor determining the propensity for an economy to experience panics.

Bank capital is understood to affect both a bank's risk-management incentives—by providing "skin in the game"—and its ability to withstand economic shocks. Thakor (2014) finds agreement in the literature that lower capital leads to higher systemic risk, but finds disagreement on the benefits of higher capital requirements in terms of reducing systemic risk versus the induced costs that may drive financial intermediation services to less regulated sectors.

Merton & Thakor (2018) distinguish bank customers from bank investors and consider how contracts between financial intermediaries and their stakeholders are structured to share risks between them. They illustrate that intermediaries have an incentive to design contracts that protect customers from credit risk; but to the extent that this insulation is costly, a trade-off must be made between the cost of protecting customers and the cost of leaving customers partially exposed. The model in Chapter 2 briefly touches on this allocation of risk between bank stakeholders, who demand safe liabilities, and those willing to bear exposure to risk through, for example, ownership of bank capital.

Key insights arising from the industrial organisation approach are that the cost of providing the services that banks provide, as well as product differentiation, and the relationships between financial intermediaries and their stakeholders, are all important factors for understanding the balance sheet structure of financial intermediaries. Chapter 2 of this thesis contributes to this discussion. The model in Chapter 2 outlines a novel paradigm for understanding the organisational structure of bank balance sheets and the nature of bank contracts. This paradigm centres on the role of banks as monetary intermediaries. As monetary intermediary, a bank's balance sheet structure, together with its demand-deposit contracts, are shown to provide an efficient and incentive compatible mechanism for transforming private debt into a cash substitute in order to mitigate various risks and costs of trading G&S. Market imperfections are shown to shape the idiosyncratic characteristics of bank balance sheets within this context. These findings differ from, but are complementary to, those of both Thakor (2014) and Merton & Thakor (2018).

1.2.4. The functional perspective approach to financial intermediation

Merton (1993; 1995) and Merton & Bodie (1995; 2005) propose a "functional perspective" approach to understanding financial intermediation. According to a functional perspective, institutional form follows from economic function, and the most efficient institutional structures for fulfilling the various functions of the financial system therefore change over time and across geopolitical locations.

Merton & Bodie describe the "neoclassical economics perspective" and the "institutional perspective" as the standard approaches to financial intermediation theory. The neoclassical economics perspective is identified as a focus on economic functions, but with no focus on the institutions that perform these functions or how they change over time. Conversely, an institutional perspective assumes that institutions matter, but takes the institutional structure as given and views the objective of public policy as assisting the institutions currently in place to survive and flourish. The functional perspective draws on both the neoclassical and institutional perspectives, but takes the economic functions performed by financial institutions as given, and then seeks to discover the best institutional structure for performing those functions at a given time and a given place.[4] In this context, the "best" financial institutional structure rests on two basic premises:

(i) that financial functions are more stable than financial institutions; and

(ii) that competition will cause changes in institutional structure to evolve toward greater efficiency in the performance of financial system functions.

[4] In this sense, it has much in common with a mechanism design philosophy; see Hurwicz (1973).

Merton & Bodie (1995; 2005) highlight resource allocation as the primary economic function of the financial system. The financial system facilitates resource allocation by allowing agents to trade financial assets, and by transforming those financial assets from one form into another. From this primary function, Merton & Bodie distinguish six "core" functions provided by the financial system. These are:

(i) ways of clearing and settling payments to facilitate trade;

(ii) a mechanism for the pooling of resources and for the subdividing of shares in various enterprises;

(iii) ways to transfer economic resources through time, across borders, and among industries;

(iv) ways of managing risk;

(v) price information to help coordinate decentralised decision making in various sectors of the economy; and

(vi) ways of dealing with the incentive problems created when one party to a transaction has information that the other party does not, or when one party acts as agent for another.

Merton (1993; 1995) argues that because the basic functions of a financial system are essentially the same in all economies, a functional perspective provides a more reliable and enduring frame of reference for understanding financial institutions than either a neoclassical or an institutional perspective.

The functional perspective approach inspires the model developed in Chapter 2. The model focuses on the exchange of G&S across time. The model derives banks as the efficient *institutional forms* of mechanisms that facilitate the economic *function* of exchanging of G&S across time. I illustrate that bank balance sheet structure and contract characteristics follow from this function, as well as the types of risks and costs in the broader institutional environment in which these exchanges take place. Banks provide transaction services that facilitate these exchanges by transforming debt into money, thereby mitigating and redistributing various risks and costs associated with these exchanges, and facilitating the optimal allocation of resources.

1.2.5. Regulation, financial stability, and bank panics

This thesis considers various policy, stability, and regulatory implications arising from the theory of monetary intermediation. This section briefly introduces these issues in the context of the financial intermediation literature.

Gorton & Winton (2003) describe the Diamond & Dybvig (1983) model as the most important study on banking panics, although they argue that the theory offers an unsatisfying explanation of panics. In Diamond & Dybvig, panic corresponds to an equilibrium in which agents believe that other agents plan to withdraw their demand deposits *en masse*, however, the theory offers no insights into what causes this belief: it is a "sunspot" event. Gorton & Winton find that the industrial organisation of the banking industry (for example, issues of monitoring and asymmetric information) is a critical factor that determines the propensity for an economy to experience panics. Additionally, they discuss what appears to be an important business cycle component to the timing of panics, with evidence that panics come at, or near, business cycle peaks. From a slightly different perspective, Borio (2014) argues that banking crises occur close to or at the peak of financial cycles.[5] Borio describes the stylised empirical features of financial cycles, but notes that there is no clear consensus on the definition of the financial cycle—although Drehmann, et al. (2012) argue that financial cycles are "distinct phenomena" from business cycles.[6] Borio (2014) discusses the policy implications of financial cycles, such as the need to strengthen the macro-prudential orientation of policy, the usefulness of building up capital and liquidity buffers, and the potential need to counteract the build-up of financial imbalances (or "lean against the wind").

Gorton & Winton (2003) note that the actual historical experience of a broad spectrum of panic-type phenomena (small losses on deposits and few bank failures) is at odds with the widely held view of panics. One problem they identify is with the definition. A broad range of phenomena are described by several often interchangeable terms, such as "financial crisis", "contagion", "banking crisis", "bank run", and "banking panic". They suggest that definitions have in common the feature that a panic is a systemic event in which consumers want to convert their demand deposits into cash, which the structure of a banking contract allows, but which the structure of bank balance sheets prevents. The conversion, or attempted conversion, is the initiating event of a panic. But there is no consensus about what events should be called "panics", or the definition of what constitutes a "systemic" event.[7] Nevertheless, Gorton &

[5] The notion of a financial cycle appears to draw parallels with the underlying ideas of Minsky's (2008 [1986]) financial instability hypothesis.

[6] Borio (2014, p. 195) summarises the financial cycle as follows: "The financial cycle is best captured by the joint behaviour of credit and property prices. It is much longer, and has a much larger amplitude, than the traditional business cycle. It is closely associated with systemic banking crises, which tend to occur close to its peak. It permits the identification of the risks of future financial crises in real time and with a good lead. And it is highly dependent on the financial, monetary and real-economy policy regimes in place".

[7] See for example Hansen (2014).

Winton find that until recent bank crises around the world, almost all the empirical work on panics had been based on the US experience. And whereas theoretical work on panics has largely been motivated by the US experience, panics are not a feature of most economies that have banks. Based on the stylised facts about cross-country banking history, they conclude that banks do not appear to be inherently unstable, as panics appear rather to be a function of industrial organisation of the banking industry, rather than the presence of banks themselves. Likewise, Corbae & Levine (2018) find that the structure of the banking system is an important determinant of financial stability.

The conclusions of Gorton & Winton are orthogonal to those of Freixas & Rochet (2008), who state that bank panics were a recurring phenomenon in the US until 1934, with similar phenomena affecting England (before the establishment of a central bank) as well as other European countries. Freixas & Rochet conclude that with many countries experiencing banking crises more recently, partly as a result of financial deregulation, it appears that without regulation, bank runs and bank panics are inherent to the nature of banking. However, Gorton & Winton (2003) question the general efficacy of banking regulation. Gorton & Winton argue that the general motivation for bank regulation is to prevent banks from moral hazard behaviour;[8] yet, they state, there is a lack of compelling evidence that moral hazard is in fact a problem. They conclude that the managerial hypothesis (i.e., nonlinear attitudes towards risk by bank managers) is empirically more important in explaining risk taking by banks than the moral hazard hypothesis—which they conclude may be rejected as an explanation.

[8] See Bhattacharya, et al. (1998) as an example of this literature.

Caprio & Honohan (2015) argue that both excessive risk-taking behaviour and moral hazard issues exacerbate banking crises.

Subsequent to the 2007/8 financial crisis, there has been significant growth in the crisis literature, with emphasis on financial stability and bank regulation. Galati & Moessner (2013) provide an overview of this research. Galati & Moessner find that the crisis highlighted the lack of analytical frameworks for predicting and coping with the build-up of financial imbalances, and a lack of understanding of system-wide risk. This has emphasised the shortcomings of a purely micro-based approach to financial regulation and supervision. These shortcomings have contributed to a growing consensus that a macro-prudential approach to regulation and supervision should be adopted. Galati & Moessner explain that there is still a lack of consensus on the objectives of macro-prudential regulation, and policy debate focuses on the usage, implementation, and effectiveness of macro-prudential tools; their impact on macroeconomic outcomes; and their relationship with monetary policy. In general, macro-prudential policy aims to support financial stability (although there is no commonly shared definition of financial stability), taking into account feedback effects that the behaviour of individual financial institutions have on each other and on the economy as a whole. Galati & Moessner summarise the general goal of macro-prudential policy as being to limit the macroeconomic risks and costs of systemic financial distress. Aikman, et al. (2018) highlight current core issues financial stability. These include debate on whether capital standards have been appropriately calibrated; the need to develop quantitative frameworks for understanding the impact of macro-prudential and micro-prudential policy on the real economy and the financial system; the future of stress-testing; and the changing size, shape, location, look, and feel of the financial system (for example, the rise of "shadow banking" and other fintech innovations).

Issues concerning financial stability and the regulation of financial intermediaries have taken centre stage in the wake of the 2007/8 financial crisis. And while regulation receives only cursory discussion in Santomero (1984) just prior to the savings and loan crisis, the issues of deposit insurance, suspension of convertibility, capital requirements, and deposit interest rate ceilings all receive attention less than a decade later in Bhattacharya & Thakor (1993). Freixas & Rochet (2008) highlight that banking regulation is ideally a response to market failure, and also results from efforts to control certain activities (such as money laundering). They identify two primary market failures with regards to the banking industry. First, the fragility of banks due to the nature of their contractual obligations, and second, negative externalities associated with bank failure. Further, Greenbaum, et al. (2016, p. 355) list the following as additional rationale for regulation:

(i) protecting the safety and soundness of an individual institution;

(ii) protecting the stability of the financial system;

(iii) fostering competition;

(iv) consumer protection;

(v) credit allocation; and

(vi) monetary control.

When implementing regulations, Freixas & Rochet (2008, p.305) classify regulatory instruments used in the banking industry into six broad types:

(i) deposit interest rate ceilings;

(ii) entry, branching, network, and merger restrictions;

(iii) portfolio restrictions, including reserve requirements;

(iv) regulatory monitoring and supervision (including closure policy);

(v) deposit insurance; and

(vi) capital requirements.

The literature aimed at establishing appropriate theory and regulation for enhancing general financial stability remains far from reaching a consensus on these matters. These issues are hotly debated and more theoretical and empirical analysis is needed. This thesis contributes to this much needed analysis by focusing on how a bank's role in the monetary system affects bank balance sheets and the broader economy.

1.3. Monetary economics

Monetary economics centres on theories of money and how the use of money improves resource allocation. Banks are therefore of interest within monetary economics because their liabilities function as money. A key difference between financial intermediation theory and monetary economics is that the underlying approach of financial intermediation theory focuses on the asset function of bank liabilities as a store-of-value, whereas monetary economics focuses on the transaction function of bank liabilities as a medium-of-exchange. The use of bank liabilities as a medium-of-exchange has been of little interest within financial intermediation theory.[9] Nevertheless, the importance to macroeconomics of recognising the function of money specifically as a medium-of-exchange has been argued since at least Clower (1967; 1970)[10] and Hahn (1973).[11].

Monetary intermediation theory focuses on the use of bank liabilities as money. This section therefore briefly introduces topics within monetary economics that are relevant to this thesis. Section 1.3.1 discusses the cash-in-advance (CIA) model. This introduces the methodology used in Chapter 4. Sections 1.3.2, 1.3.3, and 1.3.4 discuss search-theoretic models, New Monetarist literature, and the mechanism design approach respectively, all of which are related to and influence the model in Chapter 2. Section 1.3.5 introduces a number of debates concerning the role of banks in the economy from a monetary theory perspective. This provides context for the contribution to these debates made in the following chapters.

1.3.1. The cash-in-advance models of money

The CIA approach to monetary theory is a popular early attempt at integrating the transactions role of money into macroeconomic models. Boianovsky (2002) provides an

[9] "[Banks] provide transactions services ..., and they provide portfolio management services ...; the concern with banks in macroeconomics centers on their role as portfolio managers" (Fama, 1980, p. 44).
[10] "[T]he conception of money implicit in most of these discussions [within the financial intermediation literature] does not permit us to distinguish analytically between 'money' and any other commodity the demand for which depends partly on its usefulness as a store of value ... [Therefore] the premises explicitly adduced by previous writers are insufficient to support any definite conclusions about monetary optimality in a world where money commodities play a significant role as exchange intermediaries" (Clower, 1970, p. 425).
[11] "The natural place to start [discussing a theory of money] is by taking the claim that money has something to do with the activity of exchange, seriously" (Hahn, 1973, p. 234).

account of the history of the CIA approach. Boianovsky finds that the CIA approach is based on the seminal contribution of Clower (1967), and was formalised into a general equilibrium setting by the works of Grandmont & Younes (1972), Lucas (1980), Svensson (1985), and Lucas & Stokey (1987). The CIA requirement changes the constraints in the utility maximisation problem of consumer households. In the CIA setup, consumers are constrained not only by a budget constraint, but also by an additional constraint (the CIA constraint) that prevents the total value of G&S purchased by economic agents from exceeding the amount of money held over from previous periods. As opposed to the money-in-the-utility-function (MIU) or overlapping-generations (OLG) approaches, the CIA setup specifically reflects the role of money as a medium-of-exchange.

As described by Boianovsky (2002), the CIA constraint introduces a lag between the payment of cash in one transaction and its expenditure in the next. This lag means there is a time friction wedge that separates the exchange of G&S one for another. It is this time friction that I am interested in in this thesis. It is what shapes the model structures in Chapters 2 and 4. The lack of synchronisation between payments and receipts motivates a role for money in the CIA approach, and provides a transactions motive for holding money across periods. The CIA approach does not, however, make clear why agents must wait until a later period to spend their money. This lack of explanation is a lacuna in the CIA literature. This is an issue I address in the model in Chapter 4. In the dynamic stochastic general equilibrium (DSGE) model in Chapter 4, G&S are produced and sold at the beginning of a period, but agents are only paid wages at the end of a period. Labour is therefore implicitly sold on the basis of a non-tradable debt-credit contract which is only settled at the end of the period. This restricts agents to spending their wages only in a following period. This makes the CIA framework a useful template and addresses the lacuna by justifying why households must wait until a later period to spend their money. The introduction of monetary intermediation by banks enables agents to convert these otherwise non-tradable debt-credit contracts into demand deposits. Demand deposits provide a medium-of-exchange substitute to cash, which relaxes the constraint imposed by the labour contracts, thereby enabling agents to bring their spending forward and introducing a credit-based monetary system into an otherwise standard DSGE framework.

In the standard CIA setup, there are no banks, and agents have no alternative to cash as the medium-of-exchange. In the model in Chapter 4, agents can choose to pay using either cash or using bank intermediation to transform debt into a cash substitute. This provides a private and endogenous payment alternative to government issued cash. Comparing the results of the bank intermediated model to the benchmark model illustrates the dynamic effects of monetary intermediation in an otherwise standard macroeconomic environment. The use of private debt as a payment option is shown to generate a positive liquidity effect, and real variables exhibit lower standard deviations. Both of these results address two important shortcomings of the benchmark real business cycle (RBC) DSGE model.

1.3.2. Search-theoretic and overlapping-generations models of money

Search theory deals with exchange environments where acquiring information about market arrangements, relative prices, and exchange ratios is costly. In these environments, the use of money as a medium-of-exchange frees up resources otherwise allocated to the search for market information. These resources can be reallocated, thereby expanding trade.

Search theory formalises and builds on the early work of Brunner & Meltzer (1971), which explores the use of money as a friction-mitigating medium-of-exchange. Diamond (1984) presents an early search-theoretic model in which fiat money is useful for coordinating transactions across spatially separated agents. In this simple model, fiat money is exogenous and neither credit nor bank intermediation are features of the model. Townsend & Wallace (1987) extend this result. They present a model of intertemporal exchange with spatially and informationally separated markets. They illustrate that circulating private debt-credit securities increases the efficiency and scope of exchange within the model. The model does not, however, involve bank intermediation. Kiyotaki & Wright (1989) model a similar search-theoretic environment, and show that real commodity goods that are more likely to be useful in future exchange interactions are likely to emerge as generally accepted media-of-exchange—even in exchanges where they are not immediately wanted. Once again, there is no banking in the model.

Samuelson (1958) represents a pioneering use of the OLG approach to show that the use of cash as a generally accepted medium-of-exchange can provide optimal intergenerational risk sharing. Samuelson's model demonstrates that a general and important property of money is that it enables agents to transact with one another when they are otherwise unable to effectively contract directly with one another (Bryant, 1980).

Azariadis, et al. (2001) and Bullard & Smith (2003a; 2003b) present models that combine a search-theoretic approach with an OLG setup, in which spatial separation and limited communication between locations create a decentralised environment which motivates monetary exchange across generations. Both Azariadis, et al. (2001) and Bullard & Smith (2003a; 2003b) show that privately issued liabilities may improve resource allocations when cash availability is limited (I expand on this result in Chapter 2). Redemption of all privately issued liabilities is costlessly enforceable, and so both cash and privately issued liabilities circulate as substitutable media-of-exchange. There is therefore no bank-like intermediation in Azariadis, et al. (2001) or Bullard & Smith (2003a). In Bullard & Smith (2003b), exchange is subject to transaction costs; the liabilities of an intermediary, which minimise these costs, circulate as a medium-of-exchange in equilibrium. This intermediary does not, however, have the traditional structure of a bank, as its liabilities are issued with long-term maturity, while its assets have short-term maturity.

The model presented in Chapter 2 has numerous features common to these models, including: temporal separation, limited cash availability, and information frictions. But the model also deals specifically with the role of mechanisms with the specific characteristics banks in mitigating these frictions. The analysis therefore complements this search-theoretic literature.

1.3.3. *New Monetarist models of money*

Search-theoretic models of money underpin what Williamson & Wright (2010; 2011) and Lagos, et al. (2017) term "New Monetarism"—a broadly defined body of research focusing on monetary theory, monetary policy, banking, financial intermediation, payments, and asset markets.[12] Williamson & Wright (2011, p. 89) describe the principles and practices of New Monetarism as a school of thought "committed to modeling approaches that are explicit about the frictions that make monetary exchange and related arrangements socially useful, and that capture the relationship among credit, banking, and currency transactions". The approach I take in this thesis falls within this broad description.

Williamson & Wright (2010; 2011) present a baseline New Monetarist model that represents a general form of various other model specifications found within the New Monetarist literature. The model integrates both search-theoretic and competitive market environments by having agents alternate their interactions between decentralised exchange markets subject to search and commitment frictions, and frictionless centralised markets. In centralised markets, agents can enter into deposit contracts with other bank-like agents who accept and pool both productive G&S and

---

[12] This is not the same as the reference to "New Monetarism" by Kaldor (1970). Kaldor's reference is to the (new) ideas being proposed by the *Monetarists* (of the 1960s). According to Williamson & Wright (2010; 2011), New Monetarism has little in common with Old or New Keynesians and has much in common with (Old) Monetarism. But New Monetarism is an important departure from the earlier Monetarism because it takes the role of financial intermediaries and their interactions with the central bank seriously. They list five general principles of this approach: (i) a focus on microfoundations using general equilibrium, search, and game theory; (ii) a clear and internally consistent description of the exchange process (as opposed to the price setting process) and the means by which money and related institutions help facilitate that process, implying environments with explicit frictions in the exchange process; (iii) rigorous models of financial intermediation; (iv) preference for simple models and tractability; and (v) a balance of abstraction and internal consistency, meaning as few institutions as possible are taken as primitives (Williamson & Wright, 2010, pp. 18-19).

non-productive cash in exchange for "deposit" receipts. These receipts can be exchanged directly for G&S when agents are able to credibly commit to their future obligations (i.e, in centralised markets), or they can be liquidated directly into cash when agents are unable to credibly commit (i.e, in decentralised markets). In decentralised markets, exchange is constrained by search and commitment problems. Cash is useful here because it overcomes these constraints. Cash is treated as an intrinsically worthless object (meaning that it is worthless in consumption and in production), but it is able to store nominal value (an exogenous property). The Williamson & Wright studies show that cash has positive value in equilibrium as a medium-of-exchange in decentralised markets. The bank-like agents diversify across liquid cash and illiquid G&S assets in order to maximise the expected utility of depositors by providing Diamond-Dybvig like risk sharing, as well as other efficiency gains. Both bank deposit receipts and cash therefore function as media-of-exchange in this setup, but they are not direct substitutes.[13]

Gu, et al. (2016) use a New Monetarist approach to analyse cash and credit as competing payment instruments in decentralised exchange. They develop a general equilibrium framework to study how both cash and privately issued debt-credit financial instruments interact as media-of-exchange. There is decentralised exchange in an environment that incorporates spatial and temporal separation frictions, and imperfect information and limited commitment. Gu, et al. (2016) find that if credit is "easy" to obtain (meaning without constraints or restrictions), then money is irrelevant as a medium-of-exchange, whereas if credit is "tight", money is essential, but credit is

[13] Williamson & Wright (2011, p. 75) acknowledge this limitation, stating: "The model does not capture all important features of banks, such as the fact that they issue liabilities that compete with government currency in transactions".

irrelevant. The model does not, however, include bank-like intermediation: there is no conversion of debt-credit contracts into cash-on-demand-and-at-par deposits. Hence there is no specific focus on the liabilities of a bank-like entity functioning as a medium-of-exchange.

Sissoko (2016) uses a New Monetarist framework to model 19th century banking practices. Agents buy and sell at different points in time and agents are anonymous. This creates a commitment constraint, which restricts the use of bilateral credit as a means of exchange across time, and introduces a liquidity friction. Banks have a technology that enables them to keep accurate records and to address problems of anonymity. This provides a role for banks as intermediaries of general economic exchange transactions. Banks grant credit and issue notes that function as a generally accepted medium-of-exchange that responds to the needs of trade. Bank liabilities are the primary medium-of-exchange and fiat money is a subordinate substitute. The role of cash in this context is primarily to support the banking system when in crisis. This is a different starting point to typical monetary theory, as typical monetary theory does not treat bank liabilities as the *de facto* medium-of-exchange. Also, unlike typical monetary theory models, non-bank agents in Sissoko's model hold cash from period to period *because* it is interest free. Holding cash reduces the costs of otherwise entering into debt-credit contracts in order to trade. The return on cash is therefore given by the interest that is saved by reducing reliance on bank credit. Sissoko demonstrates that while this debt-credit contract arrangement meet the needs of trade, the introduction of cash may improve welfare by reducing interest rates and thus the costs of credit provision. There are two important implications of Sissoko's framework. The first is that bank lending is the key determinant of money supply. The second is that using bank liabilities as the circulating medium-of-exchange is not a matter of lending cash to borrowers, but rather a matter of using financial instruments backed by private debt as the generally accepted medium-of-exchange and a substitute to cash.

My focus in this thesis is on the use of debt-credit contracts as a medium for exchanging G&S across time, and on the role of banks in reducing the transaction costs of using these debt-credit contracts as a medium-of-exchange. Unlike Gu, et al. (2016), my focus is on the role of bank-like intermediaries transforming debt-credit contracts into money. And unlike Williamson & Wright (2010; 2011), I *do* focus on the function of bank-like liabilities as a *direct* substitute for cash. Like Sissoko (2016), I assume that transactions are fundamentally financed by unsecured debt-credit contracts driven by the "needs of trade". The model in Chapter 2 focuses on exploring the transformation of these risky debt-credit contracts into a riskless medium-of-exchange cash substitute through bank-like intermediation. The setup I use differs from Sissoko (2016) in that banks arise to deal primarily with issues of risk, uncertainty, and "bad luck", as opposed to issues of "bad faith". This thesis is therefore aligned with the New Monetarist literature, but it provides a new angle of focus.

1.3.4. Mechanism design

Hurwicz & Reiter (2006, pp. 14-15) describe an economic mechanism as follows:

> "A (decentralized) mechanism is a formal entity intended to represent a system for organizing and coordinating economic activity ... . Economic activity takes place in a setting of institutions and arrangements, a framework of laws, customs, formal

> organizations, and less formal structures. These range from relatively simple, informal, often subtle, arrangements, to highly complex and formalized structures. We refer to structures of this kind formally as mechanisms."

Wallace (2011) describes the mechanism design approach to monetary economics as a search for settings or environments in which money is an *essential* mechanism for achieving some desirable resource allocations, and in which there is no other way to achieve those allocations but through monetary trade. Wallace describes it as an approach aimed at integrating monetary economics with economic theory concerning frictions. The mechanism design approach therefore shares similar aims to the New Monetarist literature.

Kahn & Roberds (2007) use a mechanism design approach to develop a model that uses transferable debt as a medium-of-exchange. Contracts are enforced through a centralised court system, but enforcement is limited and costly. Kahn & Roberds show that in an environment characterised by uncertainty and limited enforcement, readily transferable debt (that can be used to settle other debt obligations with finality) has two advantages over simple chains of bilateral credit as a payment technology. First, transferable debt "allows for removal of less-than-perfectly reliable agents from the [credit] chain in a timely fashion", and second, "it allows agents to direct payments to the proper party without direct communication with other members of the credit chain" (Kahn & Roberds, 2007). They demonstrate that this provides useful information that improves welfare by expanding the set of efficient exchanges.

Mattesini, et al. (2009) use a mechanism design approach to study banks without making prior assumptions about what or who banks are, or what they do. They specify an environment with limited commitment, and show that some agents are better suited to the responsibility of accepting and investing deposits on behalf of others when they possess certain favourable characteristics. These characteristics include: patience; ease of being monitored; having a greater stake in or connection to the economic system; having preferential access to relatively good investment opportunities; and deriving lower payoffs from liquidating investments for selfish strategic reasons. Having these characteristics makes agents more trustworthy, meaning they are less inclined to renege on future obligations. Other economic agents are therefore more willing to extend credit against collateral that has been deposited with these trustworthy agents. Transferable claims against collateral deposited with these agents therefore enable transactions that improve the set of feasible allocations. Thus, claims against goods deposited with these trusted third-party agents can facilitate other transactions. Mattesini, et al. therefore find that trustworthiness is an important feature of banking.

In Chapter 2 I use a mechanism design approach to develop a model of monetary intermediation from first principles. I show that monetary intermediation is an efficient mechanism for allocating resources across agents with *different time preferences* in exchange environments characterised by risk and uncertainty. I show that monetary intermediation provides transaction services that mitigate and redistribute the costs of risks and uncertainties thereby improving the efficiency of trade. Importantly, I show that monetary intermediation mechanisms have distinctly bank-like characteristics.

1.3.5. *A brief review of debate concerning the role of banks within monetary economics*

Surprisingly, there are many fundamental issues related to money and banking that remain contentious within monetary economics. Arnon (2011) provides a historical account of the evolution of monetary theory that highlights these key debates. The theory presented within this thesis touches on a number of them. These include:

(i) whether money supply should be treated as primarily an exogenous phenomenon or an endogenous phenomenon;

(ii) whether banks should be restricted to being "narrow" or should be allowed to be "broad";

(iii) why certain things (for example, gold and other commodities, notes and coins, or bank liabilities) function as money and not others; and

(iv) whether banks are "creators" of money, or simply lenders of other agents' money.

These issues are in fact interrelated by a common underlying question concerning what is, or what should be, the role of the state versus private entities in control over the monetary system. The following discussion illustrates this common underlying thread.

Moore (1988) provides an in-depth discussion on whether money supply should be treated as an exogenous or an endogenous phenomenon. He meticulously unpacks how exogenous and endogenous theories of money differ. Moore highlights that the key distinguishing feature differentiating the economics of exogenous theories of money from endogenous theories, is that exogenous theories treat money supply as directly controlled and determined primarily by the state, whereas endogenous theories treat money supply as by and large a private sector phenomenon.[14]

Similarly opposing views about the role of the state in the supply of money separate proponents of narrow banking from proponents of broad banking policies. As described by Goodhart & Jensen (2015), proponents of narrow banking policies (the Currency School for example) advocate for a level or a growth rate of the aggregate money-stock that is government controlled and strictly determined by policy rules. In addition, narrow banking advocates that bank liabilities be matched by assets consisting almost exclusively of cash and other short-term government debt. By contrast, opponents of narrow banking (the Banking School for example) regard flexible, bank-intermediated, private credit as the key component of the money system. From this perspective, it is crucial for money supply to respond flexibly to the "needs of trade", and this is best achieved through bank-intermediated private credit. In this case, bank assets must necessarily include private debt. This debate has re-ignited following the financial crisis of 2007/8, and has found fuel in the discussions concerning the regulation and control of crypto-currencies (Barrdear & Kumhof, 2016; Löber & Houben, 2018). In June 2018, Switzerland held a referendum on whether their banking system should be transformed into a narrow banking system (Dawnay, 2017; Atkins, 2018). The referendum result saw a rejection of the proposal (Revill, 2018). However, the fact that the referendum took place highlights that the debate is still alive and well. The discussion in this thesis contributes directly to this debate.

[14] Wray (2012) notes that the schools of economic thought considered orthodox, such as the Monetarist, Neoclassical, and New Keynesian schools, tend to be associated with exogenous money supply theories. Whereas endogenous theories are usually associated with more heterodox schools of economic theory, such as Post-Keynesianism, or Modern Money Theory.

Theories of why certain media function as money, while others do not, are also distinguished by opposing views on the role of the state versus the role of private institutions in determining what constitutes money (Goodhart, 1998). Goodhart (1998) divides these views into three broad categories and this thesis contains elements from across all three views:

(i) Cartalism, which argues that the legal authority of the state plays the central role in determining what functions as the generally accepted medium-of-exchange through legal fiat. This general acceptability is enforced through control of what is accepted in payment of taxes;

(ii) Metallism, which argues that the use of money is primarily dependant on the intrinsic value of some real commodity. But only those commodities that have properties which are convenient for serving the function of a generally accepted medium-of-exchange (i.e., divisibility, recognisability, durability etc.) will serve as money. Metallism argues that the acceptance of this real commodity is largely independent of the state, and that the intrinsic value of the commodity is what ultimately backs the value of the money-stock; and

(iii) an eclectic body of literature, which argues that the use of anything as a generally accepted medium-of-exchange is primarily a market response to the needs of private commerce, and that this process is largely independent of the state.

This thesis supports the notion that the use of bank liabilities as a medium-of-exchange is a market response to the desires of private commerce, but the discussion does not suggest that this phenomenon is independent of the state or the broader institutional and regulatory environment. The value of a bank's liabilities, and therefore the value of the money-stock, is backed by the value of a bank's assets. Like Borio (2019), I recognise that these assets are not commodities but debt, and that a strong institutional and regulatory infrastructure is required to enforce and ensure that these obligations are honoured and are therefore of value. Within this context, I argue that the primary function of a bank is to transform private debt into money according to the private needs of trade. Like Schumpeter (1956), I recognise that underlying all trade is an exchange of G&S, and that money is the medium by which they are exchanged. The value of a bank's assets (and hence its liabilities), therefore, resides in the claim against the G&S which they represent—and the value of this claim relies on the broader institutional infrastructure that enforces debt obligations. In the following chapters, I highlight various factors worth considering when attempting to regulate banks given these considerations.

Central to the monetary system is the role of bank liabilities as money. Equally central to this system is the process by which banks "create" these liabilities. While general textbook explanations support the financial intermediation theory of the money multiplier process,[15] banks explicitly provide no intermediation services whatsoever in a growing number of contemporary descriptions of endogenous money creation (de Soto, 2009; Benes & Kumhof, 2012; McLeay, et al., 2014a; 2014b; Werner, 2014; Jakab & Kumhof, 2015). In Chapters 2 and 3, I interrogate this money creation process. I argue that the view of banks as creators of money *ex nihilo* is misleading. I argue instead that

[15] The theory of this process is increasingly criticised in mainstream literature; see for example Bindseil (2004) and Goodhart (2009).

banks transform debt into money in the form of a cash substitute.[16] I argue that this does not constitute a money multiplier process, nor does it amount to the creation of money out of nothing. The discussion on these matters draws heavily from early bank theory as described in the following section.

1.4. Early bank theory

A view of banks that is largely and glaringly absent from contemporary bank theory can be found in the early bank theory literature of Thornton (1965 [1802]), Macleod (1866; 1906), Wicksell (1936 [1898]), von Mises (1953 [1912]), Innes (1913; 1914), Hahn (2015 [1920]), Keynes (1930), and Schumpeter (1983 [1934]), among others. This early literature emphasises three points that are also central to the theory developed in this thesis:

(i) that the business of banking arises from the need for buyers and sellers to transact on the basis of debt-credit agreements;

(ii) that money is, at its core, a financial instrument arising from debt-credit relations; and

(iii) that the primary role of banks is to transform private debt-credit contracts into money.

A diverse selection of illustrative extracts from within this early literature highlights these three points.

[16] I make no comment on *why* state issued cash functions as money, but readers can see Zarlenga (2002) or Borio (2019) for different perspectives on this issue. Section 2.2 provides a brief historical account on *how* state issued cash likely came to be used as a medium-of-exchange.

1.4.1. Money: a debt-credit instrument

Early bank theory literature emphasises that money is, at its core, a financial instrument arising from debt-credit relations by which G&S are exchanged. Thus von Mises (1953 [1912]) and Schumpeter (1956) write:

> "The simple statement, [is] that money is a commodity whose economic function is to facilitate the interchange of goods and services … This applies in the first place to the function fulfilled by money *in facilitating credit transactions*. It is simplest to regard this as part of its function as medium of exchange. Credit transactions are in fact nothing but the exchange of present goods against future goods." (von Mises, 1953 [1912], pp. 34-35)

> "The monetary circulation is, in its nature and main function … nothing but a clearing system … Money is a 'claim ticket and receipt voucher' … only commodities are exchanged against commodities in the economy; money is always only an intermediate link without independent significance." (Schumpeter, 1956, pp. 155-171)

Macleod (1866, pp. 15-19) provides a lucid explanation of this viewpoint writing:

> "So long as the things exchanged were equal in value there would be no need for money. … But it would often happen that when one man required the services of his neighbour, that neighbour would not require an equal amount of service at the same time, or even perhaps any at all. If then such a transaction took place between them, with such an *unequal* result, there would remain … a DEBT—that is to say, a Right or Property would be created in the person of the creditor to demand something at some future time from the debtor … And if any man can render services to his neighbours, he must in return receive either other services, or the evidence of their being due; and if he renders more services than he immediately requires in return, he will accumulate a store of this evidence for his future wants. These simple considerations at once shew [sic] the fundamental nature of a currency … Now, as money is always exchangeable throughout the commercial community, it is evident that it has a general and a permanent value; that is, it is exchangeable among *all* persons and at *all* times, and in

> *all* places of the same country. It is therefore the highest and most general form of Credit."

And Innes (1913) writes simply and succinctly:

> "Money, then, is credit and nothing but credit. A's money is B's debt to him, and when B pays his debt, A's money disappears" (Innes, 1913, p. 402)

1.4.2. Banks as intermediaries between buyers and sellers

Unlike financial intermediation theory, early bank theory focuses on the role of banks as intermediaries between buyers and sellers exchanging G&S on the basis of debt-credit contracts. As Innes (1913) explains:

> "Debts and credits are perpetually trying to get into touch with one another, so that they may be written off against each other, and it is the business of the banker to bring them together … either by *discounting bills*, or by *making loans*. … The process of discounting bills is as follows: A sells goods to B, C and D, who thereby become A's debtors and give him their acknowledgments of indebtedness, which are technically called *bills of exchange* … . The banker buys from A the bills held by him on B, C and D, and A now becomes the creditor of the banker, the latter in his turn becoming the creditor of B, C and D … [T]he same result is reached by means of a loan … . In this case the banking operation, instead of following the sale and purchase, anticipates it. B, C and D before buying the goods they require make an agreement with the banker by which he undertakes to become the debtor of A in their place, while they at the same time agree to become the debtors of the banker. Having made this agreement B, C and D make their purchases from A and instead of giving him their bills which he sells to the banker, they give him a bill direct on the banker." (Innes, 1913, pp. 402-403)

Likewise, Macleod (1866) writes:

> "The immense majority of commercial bills are not paid in money, but by credits in bankers' books. Most men in commerce draw bills and accept bills, that is, they have debts due to them, and debts due by them. These fall due at different dates, and when a trader's acceptances are falling due, he takes some of the debts due to him to his banker, and sells them to him. The banker buys them … by creating fresh liabilities of his own, and writing down so many figures to his customer's credit. When his [the trader's] own acceptance falls due, and is presented to him for payment, he draws a cheque upon his banker, and if the holder of the bill is a customer of the same banker, the matter is settled by a mere transfer of figures in the banker's books; if he [the holder of the trader's bill] is the customer of another bank, the two bankers have probably an exchange of debts, arising out of similar transactions on both their parts" (Macleod, 1866, pp. 125-126)

And Wicksell (1936 [1898]) writes:

> "The actual exchange of commodities proceeds very simply. The buyer draws a cheque … for the appropriate sum, and the seller cashes the cheque, the sum being thus credited to him by the Bank … During this time, the sellers are in reality extending credit to the buyers to the amount of the sum in question … although on the surface the payment has the appearance of being immediate. This is brought about as a result of the facilities and the guarantees provided by the Bank." (Wicksell, 1936 [1898], pp. 71-72)

1.4.3. Transforming debt into money

Early bank literature highlights the role of bank liabilities as a payment alternative to cash. Within the context of banking practice in these earlier periods, this had long been understood; as de Roover (1948, p. 283) writes:

> “As early as 1421 a Venetian document distinguishes between two kinds of money and uses the expression *contadi di banco* for bank money and *denari contadi* for coin. Later one finds even the phrase *moneta di banco* which is a literal translation of ‘bank money’. The use of such expressions implies a clear recognition of the fact that bank deposits are money and perform the function of money.”

And while both coins and bank liabilities are recognised as money in early bank theory literature, it is also clear that the this literature recognises how the actual institutional form of money is influenced by the broader institutional and regulatory environment. As Macleod (1866, pp. 119-122) writes:

> “[I]n 1709, on the renewal of the charter, Parliament meant to confer a monopoly of banking on the Bank of England … And this privilege of banking … meant the creation and issuing of ‘Currency;’ and to prohibit persons from creating currency was, in fact, to prohibit them from doing banking business … It was supposed that no legal ingenuity could devise an expedient to evade so extensive a prohibition. The form of words adopted in this Act, was devised in reference to the methods of doing banking business at the time they were framed, and they were successful in preventing any rival bank being formed, so long as bankers adhered to that particular method of doing business. But about 30 years afterwards, bankers adopted a change in method of doing their business, so simple, and apparently so unimportant, as scarcely to deserve attention. And it was this mere change in the form of doing their business, that is, of creating liabilities, or currency, in a form not provided for by the words of the Act, that cut away the ground from under the Act, and was the means whereby the present Joint Stock Banks in London were formed, and thus destroyed the monopoly of the Bank of England.”

Regardless of institutional form, an important feature differentiating money from other financial instruments in early literature is what Schumpeter (1983 [1934]) refers to as “purchasing power”—which is to say that money functions as a generally accepted medium-of-exchange that converts into G&S on demand and at par. In this regard, banks are described as providing new purchasing power by transforming private debt-credit contracts into money. In early bank theory, the institutional infrastructure that enables private debt-credit contracts to embody *new* purchasing power is central to the understanding of banks. Schumpeter (1983 [1934], pp. 72-73) writes:

> "[From] whence come the sums needed to purchase the means of production necessary for the new combinations if the individual concerned does not happen to have them? ... This method of obtaining money is the creation of purchasing power by banks."

By contrast, from the perspective of financial intermediation theory, the role of banks in the provision of *new* purchasing power is of little interest. Instead, banks are a mechanism for exchanging financial instruments that do not have purchasing power (i.e., debt or equity based financial instruments), for *already existing* purchasing power in the form of cash. In terms of financial intermediation theory, banks facilitate the transfer of *already existing* purchasing power from lenders to borrowers, and they leverage this purchasing power through a multiplier effect. Banks are viewed as an alternative and competing mechanism to financial markets for allocating already existing purchasing power. The role of both banks and financial markets in financial intermediation theory is to bring together both savers and borrowers, by means of which savers transfer loanable funds with already existing purchasing power to borrowers. Through financial markets and other financial intermediaries, savers lend savings (implicitly in the form of cash) to investors in return for investments in the form of debt or equity based financial instruments (Mishkin & Eakins, 2006; Greenbaum, et al., 2016). By contrast Schumpeter (2006 [1954]) and Macleod (1866) write:

> "Banks are no longer said to 'lend their deposits' or 'other people's money,' but to 'create' deposits or bank notes ... Bank notes and checking deposits eminently do what money does; hence they are money. Thus credit instruments, or some of them, intrude into the monetary system." (Schumpeter, 2006 [1954], pp. 304-5)

> "[There is a] very great misconception that is likely to be produced by an expression which is very often used regarding bankers, that they are merely agents between persons who want to lend, and those who want to borrow. This is not true in the ordinary sense of the words lending and borrowing ... The common notion of banking is, that it consists in lending money upon the security of bills of exchange. Such an idea, is profoundly erroneous, as it [instead] consists in buying debts with 'promises to pay,' or creating liabilities ... [T]he cheques drawn against these credits, so created by the banker, circulate commodities exactly in the same manner as bank notes do, which circulate commodities exactly in the same manner that gold and silver money does. Consequently, these credits, so created by the banker, are currency, or circulating medium." (Macleod, 1866, pp. 127-128)

Given that modern banking theory discourse takes place primarily from the portfolio management perspective of financial intermediation theory, it is clear that this "Liquidity Theory of Money" focus has shifted away from the transactions perspective of earlier bank literature.[1]

1.4.4. A modern interpretation of early bank theory

A modern interpretation of this early literature is provided by Sissoko (2015a; 2015b; 2016). Sissoko describes and models the basic elements of early bank theory, arguing that this early theory accompanied the growth of modern banking and is therefore a valuable framework for: (i) understanding the development of the modern financial

[1] Lucas (1984) offers an explanation for this phenomenon.

system; (ii) understanding the relationship between banks, financial markets, and the central bank; (iii) evaluating the causes of the recent financial crisis; and (iv) evaluating theories of macro-prudential regulation and reform. The theory presented in this thesis is in broad agreement with these statements.

Sissoko identifies two important principles underlying the business of banking in the 19th century which differ from traditional economic models. The first is that optimal money supply is driven by the unpredictable "needs of trade" and is inherently stochastic. The second is that transactions are fundamentally financed through unsecured debt-credit contracts. The role of the banking system in this early literature is the monetisation of these transaction-based debt-credit contracts. Banks are understood to facilitate trade between buyers and sellers of G&S by accepting the debt obligations of buyers as assets on their balance sheets, and matching these assets with corresponding demand-deposit liabilities for the sellers.[2] As a result, a key feature of Sissoko's (2016) banking framework, is based on an underlying notion that money supply is primarily a private sector phenomenon, and that the role of government fiat money (cash) is fundamentally to support and stabilise the predominantly privately issued (bank-money) money supply.

Sissoko uses this framework to present a bank-centred view of the money market in which all demand and other short-term bank liabilities (including contingent liabilities) are treated as potential money, and banks are acknowledged as different from

[2] As Sissoko (2016, p. 2) writes: "when banking is studied using the new monetarist framework, one of two approaches is typically taken: either banks are modeled as a Diamond-Dybvig's mutual-fund-like intermediaries that receives [sic] deposits of real goods and then invest those goods on behalf of their depositors or note-holders[,] … or banks are agents with superior commitment ability who issue inside money … [By contrast] Classical banking theorists view the key banking activity not as the investment of deposits, but as the exchange of two types of debt: for a fee a bank is willing to allow a buyer to trade using the bank's liability, and accepts in exchange the liability of the buyer to the bank."

other financial intermediaries *because* their liabilities circulate as money. From this perspective, shadow banking represents a shift in the type of liabilities that support the money supply.

The discussions in this thesis draw strong parallels to the studies of Sissoko, but the setup differs. As a result, a number of the policy conclusions drawn in this thesis contrast to those of Sissoko. For example, I do not conclude that maturity transformation by banks should be restricted to short-term financing (this can be interpreted Sissoko's version of Adam Smith's Real Bills Doctrine). On the contrary, I argue that both the long- and short-term financing activities of banks can be used as policy tools for enhancing financial stability (see Chapter 3). Unlike Sissoko, I do not conclude that the integration of money and capital markets necessarily promotes financial instability. Instead, I argue that there are complementarities that arise from the integration of money and capital markets (see Chapters 2 and 3). Like Sissoko, this thesis represents a modern revival of early bank theory expressed by the likes of Wicksell and Schumpeter.

1.5. Concluding remarks

Why do banks exist? This question is at the heart of bank theory. In this thesis I revive, develop, and expand on an old answer to this fundamental question. I present a theory of banks as monetary intermediaries. I show that as monetary intermediaries, the *raison d'être* of banks is to mitigate the uncertainties, risks, and other frictions associated with the exchange of G&S backwards and forwards through time. As *monetary* intermediaries, banks mitigate these transaction costs by transforming private debt-credit contracts into a fungible cash substitute. This facilitates the intertemporal exchange of G&S. These transaction costs are mitigated, reduced, and redistributed through a bank's balance sheet composition and structure.

This chapter situates the contributions of this thesis within the context of the broader body of bank theory, and lays a contextual foundation for the material presented in subsequent chapters. The theory developed in this thesis provides insight into various questions, puzzles, and lacunae within bank theory that are highlighted within the discussion of this chapter. These include: why run-able bank demand deposits are used to provide liquidity and transaction services, why bank intermediation is more pervasive in more informationally complete economies, and how the use of bank liabilities as a medium-of-exchange improves resource allocation. In the following chapters, I provide novel answers to these questions, and contribute towards the growing body of research focusing on the role of banks as friction-mitigating intermediaries whose liabilities function as a generally accepted medium-of-exchange.

Chapter 2

Monetary Intermediation Theory

"Thus we see that the distinctive function of a bank, and a banker, was to issue notes, payable to bearer on demand, which were to circulate as money. *That is to create paper currency* ... And the meaning of the word 'to bank' was to buy money and bills of exchange with such notes ..."

— Macleod, (1866, p. 118)

Part 1: The Role of Banks as Monetary Intermediaries

2.1. Introduction

The use of bank liabilities as a medium-of-exchange is unique to the banking firm. But why do the liabilities of banks perform this function? The liabilities of other firms do not serve this monetary role. This chapter argues that bank liabilities function as money because banks are designed to be *monetary* intermediaries. As monetary intermediaries, banks provide services that facilitate transactions between buyers and sellers by transforming risky debt-credit contracts (bank assets) into safe money (bank liabilities). In this chapter, I develop the concept of monetary intermediation and model the role of banks as monetary intermediaries. This advances the theory of banks by developing a framing of a bank's intermediary role in the economy based on the transaction services they provide.

Establishing the role of banks as monetary intermediaries is important for at least three reasons. First, establishing the role of banks as monetary intermediaries provides a new explanation for the characteristics and organisational structure of banks. This provides answers to a number of empirical puzzles within bank theory (see Part 2 of this chapter). Second, the *framing* of a bank's role in the monetary system influences and informs understanding and public opinion about a bank's role in the monetary system and in the economy in general. This in turn influences the debate that ultimately shapes policy and regulation (see Chapter 3). The third reason why establishing the role of banks as monetary intermediaries is relevant, is the urgent need to more effectively incorporate the role of banks into macroeconomic models after the 2007/8 financial crisis. As Jakab & Kumhof (2015) illustrate, the dynamic effects of incorporating banks into otherwise standard macroeconomic models depend crucially on how a bank's role in the monetary system is specified. In Chapter 4, I show how incorporating the role of banks as monetary intermediaries into a standard DSGE framework affects the outcomes of these models.

The first part of this chapter formally introduces and describes a bank's role as monetary intermediary. It is structured as follows. Section 2.2 traces the historical context and intuition underlying the theory of monetary intermediation—which is centred on the use of debt-credit agreements as media-of-exchange. Section 2.3 outlines the process of a debt-credit financed transaction without the intermediation services of a bank. Section 2.4 briefly explains the emergence of modern banks as intermediaries to these transactions. Section 2.5 highlights the risk and cost-mitigating function of banks as intermediaries to these transactions. Section 2.6 discusses how the risk- and cost-mitigating function of banks determines the use of bank liabilities as a cash substitute. Section 2.7 considers the implications of the discussion on various issues related to the role of banks in the monetary system. Section 2.8 concludes the first part of this chapter.

2.2. A brief history of debt and cash as media-of-exchange

Tracing the emergence of debt and cash as media-of-exchange provides insightful historical context to the content of this thesis. And as Ingham (2000, p. 27) writes:

> "The use of specific institutionally-legitimate debts as means of payment is arguably one of the most important developments in the history of humanity's organizational or infrastructural power."

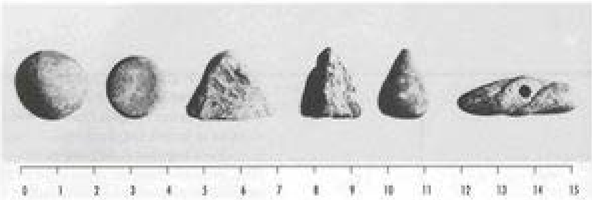
The earliest records of private debt date back roughly 10,000 years to ancient Mesopotamia in around 8,000 B.C.E. (Schmandt-Besserat, 1980; 1991). By this stage of society's development, the concepts of owing and being owed were well established within communal settings. More formal acceptance of individual property rights began reflecting in the emergence of private seals for marking ownership of property, and of private debt-credit contracts as a means of exchanging property (Schmandt-Besserat, 1980; Duistermaat, 2010; Frangipane, 2016). These contracts (recorded thousands of years before the emergence of writing) were expressed in the form of three-dimensional baked clay shapes, or tokens, representing the types and quantities of the goods being recorded (see Figure 2-1 (i) below). These tokens were wrapped in hollow spherical clay "envelopes", which were sealed and baked so that the terms of the contract were immutable (see Figure 2-1 (ii)). The limited variety of tokens in these early contracts corresponded to a limited variety of goods produced and exchanged during this early period (Schmandt-Besserat, 1980; 1991).

Over the next 5,000 years, advances in agricultural technologies saw the emergence of more intense agriculture and more sedentary communities. These changes were accompanied by increasing surpluses that provided for an increasing variety of G&S (reflected in an increased variety of contract tokens; see Figure 2-1 (iii)), and increasingly complex societies now administered by temple-based bureaucrats (Schmandt-Besserat, 1980; 1991). These temples became central to the socio-economic activities and well-being of these societies (Schmandt-Besserat, 1991; Semenova, 2011; Frangipane, 2016), including the administration and governance of both private and public, inter-personal and spiritual, debt-credit relations (Ingham, 2000; Hudson, 2004). At a societal level, temples kept detailed records of holy tributes owed or paid to the temples using marked clay tokens.[1] In some cases, temples issued stamped *metallic* tokens in acknowledgement of tributes received. These were ceremoniously returned to the temple to partake in important and sacred feasting and prostitution rituals aimed at promoting fertility and fruitful harvests (Kurke, 1999; Lietaer, 2002; Semenova, 2011). These, and other similar temple-oriented practices, appear to have been early forerunners of coins issued by later state institutions, with tributes replaced by state purchases, and religious ceremony replaced by (certainly less titillating) taxes and government service provision (Schmandt-Besserat, 1980; Semenova, 2011).[2] However, the exact details of the evolution of this system to one of formal coinage and taxation remains unclear (Zarlenga, 2002; Graeber, 2011); it was another 2,000 years before the first obviously state issued coins were minted in Lydia in about the 7th century B.C.E (Gardner, 1918; Kagan, 1982). Egypt began issuing coins from about the 3rd century B.C.E., but private debt-credit contracts continued to function as state accepted payment media alongside these coins (von Reden, 2007).

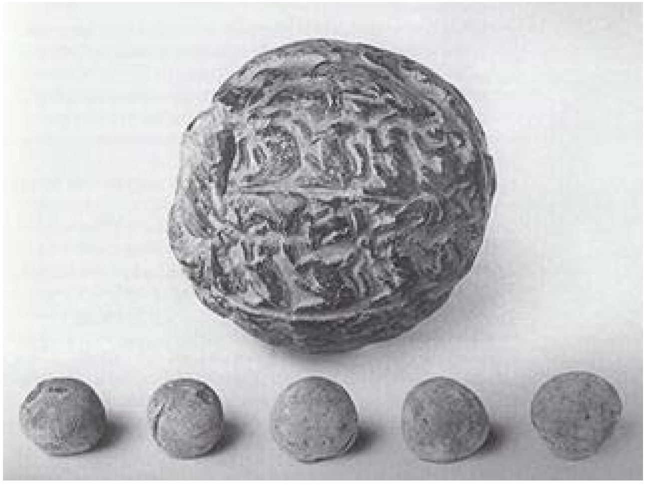
[1] This accounting system gradually evolved into the earliest forms of cuneiform writing (Schmandt-Besserat, 1980; 1991)
[2] The English word "money" originates from the name of the Roman goddess Juno Moneta, from whose temple coins were minted in ancient Rome (Lewis & Short, 1879).

Figure 2-1: Clay tokens and envelopes

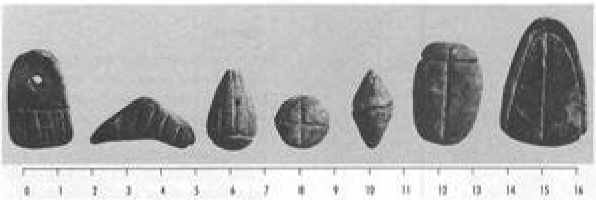
(i) Plain tokens*:



(ii) Clay envelope and contents*:



(iii) Complex tokens*:



*Source: Schmandt-Besserat (1991).

This mixture of both government and privately issued debt-credit based financial instruments of exchange continued to characterise commercial life in medieval Europe (de Roover, 1948); it characterised commerce in the industrialised economies of 19th and 20th century Europe and America (Macleod, 1866; Wicksell, 1936 [1898]; Gillet Bros. Discount Co. Ltd, 1964); and it continues to characterise today's advanced economies in the form of cash and commercial bank liabilities. Recognition of these circumstances provides historical context to the following discussion.

2.3. Debt-credit transactions without banks

To illustrate the role of banks as monetary intermediaries, it is useful to begin by considering a monetary economy in which two individuals, a seller (Ms S) and a buyer (Mr B), want to transact. B wants to buy a residential property from S.[3] S owns a residential property and wants to sell her property to B. S is willing to transfer her property to B and B is willing to accept transfer of the property. The price of the transaction is agreed upon at say $100,000. This is not a direct barter exchange transaction, but B also does not have $100,000 in cash available. Neither S nor B have access to a formal banking institution (or any other type of formal financial institution).[4] Therefore B must come to some debt-credit agreement directly with S (B may offer to pay off the $100,000 in, say, monthly cash instalments over a 20-year future period). If S agrees, S and B enter into a written contract setting out their respective rights and obligations. This contract constitutes a debt-credit contract. It is a financial instrument and is a form of promissory note.[5] Even in today's modern monetary economies, entering into similar debt-credit contracts of exchange need not directly involve a bank—as evidenced by today's shadow banking systems (Bernanke, 2012).

B's offer in exchange for the residential property specifies that B will make periodic payments to S over a 20-year period. Before the transaction, S had an asset in

[3] I use residential property in this narrative simply because residential mortgage loans are of particular contemporary interest. This is because sub-prime residential mortgage lending is placed at the heart of explanations of the 2007/8 financial crisis—see for example the textbook discussion in Mankiw (2013, pp. 575-589).

[4] Lack of access to formal financial institutions remains commonplace in low-income countries and rural economies today (Adams & Vogel, 1986; Besley, 1995). Taking this limited institutional setting as a starting point is therefore not without practical contemporary value.

[5] In this specific case, the promissory note may be in the form of a mortgage note and it may be accompanied by a Deed of Trust contract.

the form of a residential property valued at $100,000; after the transaction, S has an asset in the form of a claim against B for future periodic payments—the value of which (ignoring risk premiums, discount factors, and the time value of money for the moment), is also $100,000. This is a *quid pro quo* process. After the transaction, S has exchanged one form of property (residential property) for another (a claim for future cash payments).[6] B has an asset in the form of a $100,000 house and a matching liability obligation to make future payments. The net asset positions of both agents remain the same both before and after the transaction.

The obligation to make future payment and the corresponding right to this payment (both of which will be recorded as financial instruments in the form of, for example, bills of exchange or promissory notes) arise as a result of an underlying transaction. B may be required to pledge collateral in order to mitigate the risk of future non-payment to S, but the $-value of any collateral provided does not determine the $-price of the residential property. A financial instrument is created between the two parties that records the value exchanged, and this instrument obtains value as a direct consequence of the agreed upon value of the transaction. The financial instrument *derives* its value from the agreed upon value of the residential property,[7] but the value of the instrument is *underwritten* by the obligation of B's future payment. This is an important point that also applies to a bank's balance sheet—the value of a bank's liabilities is underwritten by the value of the bank's loan assets.

[6] B's debt to S has long been recognised as incorporeal property—see Macleod (1866).
[7] I believe this is what von Mises (1953 [1912]) refers to as the "objective exchange-value of money".

2.4. The emergence of modern banking institutions

Prior to the emergence of modern banking institutions and the development of modern financial markets, private debt-credit contracts—like the arrangement described above—were common formal methods of exchange (Kahn & Roberds, 2007).[8] The use of these financial instruments as media-of-exchange extends from early societies to recent times.[9] Macleod (1866), Wicksell (1936 [1898]), and Innes (1913) describe how circulating bills of exchange were gradually replaced by modern banking facilities throughout the 17th, 18th, and 19th centuries. Modern banking offers two advantages over the bills-of-exchange system. The first relates to the ease at which claims against the debt underwriting these instruments can be transferred (Kahn & Roberds, 2007). The second relates to the exposure of the bill holder to the idiosyncratic risks of default associated with the issuers of the bills. Prior to modern banking, the transferability (or "negotiability") and general acceptability of these bills was greatly improved through a system of concurrent endorsements of the underlying debt by successive holders of a bill—the so-called "names". As Macleod (1866, p. 95) writes:

> "Some thirty years ago, almost the entire circulating medium of Lancashire consisted of Bills of Exchange, and they sometimes had as many as 150 indorsements upon them before they came to maturity."

8 This was the case despite the presence of sophisticated financial markets, suggesting this method of financing exchange held advantages over using financial markets.

9 Graeber (2011) provides an account of the transactional role of debt over the last 5,000 years, and its link to the modern monetary system. De Roover (1948) describes these practices in medieval Europe. Sissoko (2015a; 2015b) reviews early discount banking practices of the 18th and 19th centuries and explains how the modern monetary system descends from these earlier practices. Gillet Bros. Discount Co. Ltd (1964) provides a thorough description of the practice occurring well into the second half of the 20th century. Financial columnist Frances Coppola offers a personal contemporary description of how this still applies today (Coppola, 2012).

Each endorsement remained in effect until final settlement of the bill. Each endorsement reduced the default risk exposure for each successive holder of the bill. This reduction in risk enabled these bills to circulate widely as commonly accepted media-of-exchange. Both Macleod and Wicksell detail how this convoluted process of endorsement and transfer was gradually replaced by the accounting entries in the books of a *trusted* banking intermediary.

Modern banking emerged as a convenient method of increasing both the transferability and the acceptability of these private contracts. The process was as follows. Instead of endorsing a bill and then trading it for G&S, a merchant would cede the bill—at a discount—to a well-known and reputable intermediary institution in exchange for a corresponding liability issued against the institution itself. The institution's liabilities (which were underwritten by the underlying bill) would circulate widely and could be traded for G&S as a more *generally accepted* medium-of-exchange. The final settlement of the original ceded debt was administered by the intermediary institution itself. These institutions were called banks, and their circulating liabilities were referred to as bank notes (Macleod, 1866). These bank notes have been considered money since their early emergence (de Roover, 1948). These physical bank notes have gradually (at times, due to regulation) been replaced by alternative bank products, including cheques, credit or debit cards, and electronic fund transfers—see for example the discussion by Macleod on the effects of regulation (1866, pp. 119-122).

2.5. Credit transactions with banks

Returning to the transaction between B and S, the arrangement (without the involvement of a banking intermediary) holds a variety of potential costs and risks for S. There is: (i) a time-preference liquidity constraint; (ii) solvency and default risk; and (iii) a 20-year administrative burden placed on S. S could mitigate some of this risk by carrying out a due diligence exercise and by insisting on a mortgage bond. But amongst other things, this would require that S has the skills and information necessary to conduct the due diligence exercise, and the necessary knowledge about mortgage bonds. Consistent with both the transactions theory of the firm (Coase, 1937) and historical evidence (de Roover, 1948; Gillet Bros. Discount Co. Ltd, 1964; Graeber, 2011), an intermediary can be expected to emerge if it is able to reduce the broad spectrum of transaction costs (risks, uncertainties, and other costs) associated with this transaction.[10] S's transaction costs can be mitigated by calling on a bank.[11]

Assuming that a bank can now be called upon to assist, S may instead agree to sell to B on condition that B negotiates the services of a bank to act as an intermediary to the transaction.[12] Having entered into a suspensive contract of purchase (the offer to purchase, potentially subject to a mortgage bond in this case), B approaches a bank and the bank carries out a due diligence exercise to evaluate the financial health of B. The bank, dealing with numerous such transactions, acquires the specialised skills necessary

[10] As Kahn & Roberds (2009) state: "The history of banking … indicates that banks arose first as providers of payment services … and then diversified into the business of financial intermediation".

[11] Here, a bank refers to a typical bank balance sheet structure consisting of an asset portfolio of private debt matched by demand and other deposit liabilities.

[12] Alternatively, B, through financial markets, could enter into a debt-credit agreement with some third party who has the available cash. And having thereby acquired the cash from the third party, pay this cash over to S. Financial markets therefore provide a complementary method of financing this transaction where available and where cost effective.

to assess whether B is a good default risk or not. In addition, the bank can reduce the costs of idiosyncratic default risk through diversification across many buyers, through insurance contracts, and through various forms of security (such as a mortgage bond). Through specialisation and economies of scale, banks can minimise the costs associated with entering into these contracts. If the bank is satisfied with the risk profile of B to the extent that the bank is confident that B will make the required periodic payments, the bank issues B with a guarantee to make a cash facility available on demand to S for an amount of $100,000. The guarantee is presented to S (or more likely her legal representative), and on transfer of the property, it is returned to the bank (or more likely the bank's legal representative), whereby S acquires the right to make withdrawals of cash up to $100,000 on demand from the bank. At this point, the bank, as intermediary, incurs both the responsibilities of monitoring the debt obligation of B to ensure payment, and of converting the liability to S into cash-on-demand-and-at-par when required.

For the risk- and cost-mitigating services provided by the bank to S, the bank will be compensated from an interest cost (or fees charged) built into the payments of B. B benefits from the intermediation of the bank if the sale would not have been conducted without the services offered by the bank, and B potentially incurs a lower cost than he would otherwise be afforded without the specialist skills and services of the bank (Tobin, 1963).

2.6. Bank-money

The bank alleviates S's liquidity constraint by providing liquidity services to S that guarantee immediate access to cash-on-demand-and-at-par. The bank makes it clear to S that she can make withdrawals from the bank's coffers to the value of the transaction at any time. Being a claim for cash-on-demand, S's asset is no longer a claim for payment from B, but a claim for cash on-demand-and-at-par for an amount of $100,000 from the bank. This obligation is recorded by the bank as a liability accounting entry called a "demand" deposit. The demand-deposit contract is such that this liability obligation is easily transferable at minimal cost. As a result of this guarantee of convertibility, and the safety and ease of transferability, demand deposits additionally provide transaction services by functioning *directly* as a medium-of-exchange substitute to cash without requiring conversion to cash (Macleod, 1866; Gorton & Winton, 2003). This reduces the amount of cash that the bank might otherwise need to hold in its coffers to meet its guarantee of convertibility.

The bank's liability obligation to S reflects the value of an underlying transaction which takes place externally to the bank, and the value of which is *recorded* by the bank. The claim against B is the bank's matching asset. With this arrangement, S has resolved her liquidity, default, and administrative problems. S is now exposed to both the risk of the bank's liquidity and its solvency. To prevent bankruptcy, the reputation and management of the bank as well as legislation and regulation must be such that S, and indeed others, have confidence that the bank will provide cash-on-demand when requested. When these conditions are in place, the liabilities of banks function as a generally accepted medium-of-exchange and as substitutes for cash. The position of the bank on conclusion of the sale is illustrated in Figure 2-2 below.[13] If confidence in the bank's ability to honour its obligations wanes, S and the bank's other demand-deposit clients may approach the bank to make good on its obligation to convert its liabilities into cash. If the bank is unable to do so in sufficient quantity, it will face a liquidity shortage. This could result in insolvency if the bank is forced to sell its assets at a sufficiently large discount to meet its liquidity requirements. This is an important disciplining condition for the bank.

Figure 2-2: Bank balance sheet as a result of monetary intermediation

		Assets (loan)	Liabilities (deposit)
Party	Rights or responsibilities	Dr	Cr
Mr B (buyer)	Obligation to make re-payment of the loan	$100,000	
Ms S (seller)	Right to make cash withdrawal		$100,000

2.7. Implications for understanding the role of banks

Two observations can now be made concerning what has transpired. First, money supply (as measured by the sum of cash and bank deposit liabilities in circulation in the economy) has increased despite the fact that no cash has been physically deposited with the bank. Second, in what is now a social convention, the seller will put the property on the market, subject to the condition that the prospective buyer will obtain financing from a bank; that is, on condition that a bank will accept the offer of periodic payments

[13] Nuances concerning the timing of a bank's balance sheet entries are discussed in Appendix 2A.

from the buyer. To an outsider simply viewing the books of the bank, it may appear as if by recording an asset accounting entry and by recording a corresponding deposit liability entry, the bank has created money out of nothing, but this is only the *prima facie* appearance and not the truth of the matter because the outside observer has neglected to acknowledge that the deposit value records a value-for-value exchange transaction.[14] In reality, the seller no longer possesses a house and the buyer now possesses a house.[15]

These observations have at least five implications for the debate on the role of banks in the monetary system. The first is that the $100,000 of recorded bank deposits was not created *ex nihilo*. Bank demand-deposit liabilities derive their value from the value of the G&S exchanged for the debts held within the bank's asset portfolio.[16] B's indebtedness does not arise out of nothing, but in exchange for G&S received in the present and for which payment is expected to be received (and will be collected and administered by the bank) in the future (Schumpeter, 1983 [1934]). A financial instrument (bank-money) with a face value of $100,000 was created as a result of an intertemporal value-for-value exchange: the exchange of $100,000 worth of property.

Second, the bank as monetary intermediary acquired a claim against B for $100,000 and a liability towards S for the same $100,000. The bank therein acts as intermediary to the underlying exchange transaction. The bank's liability to S provides S with a cash substitute, which reduces the various risks and other transaction costs faced by S. The bank must manage and monitor the debt obligation of B to ensure that the

[14] See Chapter 3 for a more in-depth discussion on this point.
[15] The ownership of the house by the buyer will likely be subject to a mortgage bond in this case.
[16] How the creation of bank-money affects general price levels or the value of the bank-money itself is an important issue, but is beyond the scope of this particular discussion.

value of this debt is not undermined, and to ensure that the bank itself remains a solvent going concern. To manage its liquidity position, the bank must monitor the net demand for cash by its depositors as well as transfers to and from competitor banks. The rationale for the existence of the bank in this process is the reduction of the various risks and costs associated with the underlying exchange transaction. In the process, the bank has become exposed to both the default and solvency risks of the buyer, and the liquidity risks associated with the potential cash conversion demands of the seller.[17] These are core aspects of banking. It is the bank's ability and skill to manage these risks and to take over the 20-year administration of the debt-credit agreement which justify its existence.

The third implication is that the quality of a bank's deposits is underwritten by a portfolio of private debt obligations. The bank's deposits are a significant component of the money supply, and hence a bank's lending decisions impact on the quality of the money supply. If the quality of the assets on bank balance sheets is undermined, the value of the bank-money stock is undermined. Given that bank guarantees of conversion to cash-on-demand-and-at-par are underwritten by these assets, a negative shock to the value of a bank's asset portfolio puts pressure on both bank liquidity and bank solvency. This has the potential to undermine the monetary system. Problems in the monetary system lead to strains on the economic system and influence broader financial stability. It is therefore imperative for maintaining the quality of the circulating medium-of-exchange, and for overall financial stability, that the quality of

[17] Should the seller wish to convert the bank's demand-deposit liability (i.e., the seller's bank-money) into cash-money as guaranteed, it is the obligation of the bank to provide the cash-money. This aspect of banking is discussed further in Part 2 of this chapter.

private debt underwriting bank deposits is of the highest pedigree (Borio, 2014; Mian & Sufi, 2015).

Fourth, bank loans and deposits are the result of general economic transactions. One could therefore reasonably expect bank-money and bank intermediation to be positively correlated with gross domestic product (GDP) as found in King & Levine (1993) and Levine & Zervos (1998). What is more, one would expect monetary intermediation to be a boon to economic activity.

The fifth implication is that bank loans are not dependent on prior savings.[18] In this regard, the role of banks in the money creation process is not a function of savings but a function instead of:

(i) the desire of economic agents to exchange G&S on the basis of private debt-credit contracts;[19]

(ii) the desire of economic agents to hold bank deposit liabilities as a medium-of-exchange;[20]

(iii) the bank's administrative expertise and ability to manage and reduce the risks and costs related to loan default and deposit withdrawal;[21] and

(iv) the bank and its stakeholders' expectations regarding the bank's profitability and solvency.[22]

[18] Unlike raising financing through financial markets.
[19] See Macleod (1866; 1889).
[20] This corresponds to the discussion in Gorton (2014) and Sissoko (2016).
[21] This corresponds to the discussion in Tobin (1963).
[22] See Jakab & Kumhof (2015). Thornton (1965 [1802], p. 77) writes: "The apprehension of loss, is, therefore, continually operating on the mind of the lender as a restraint on the custom of giving credit, while the compensation he receives for the use of the capital which he supplies, acts as an encouragement to the practice".

The salient features of this discussion extend to all manner of intertemporal exchange transactions for which buyers cannot immediately pay with money, and which are thus facilitated by means of a bank. This explanation is congruent with early banking theory literature, it complements financial intermediation theory, it complements monetary theory, and it complements the discussion related to theories of endogenous money creation *ex nihilo*.

2.8. Conclusions

Bank liabilities function as money because banks are monetary intermediaries. Monetary intermediation not only moves resources forward and backward through time, it reduces the risks and costs of doing so. As monetary intermediaries, the role of banks is to facilitate transactions between buyers and sellers by transforming risky, private debt-credit contracts into a safe cash substitute. This mitigates the various risks and other transaction costs faced by buyers and sellers exchanging G&S on the basis of these debt-credit contracts. The salient features of a bank's monetary intermediation role can be summarised as follows:

(i) a bank's balance sheet records the nominal value of transactions intermediated by the bank;

(ii) these transactions are between buyers and sellers of G&S generally;

(iii) the bank provides credit services to the buyers and administers the risks and costs of the buyers' debt obligations;

(iv) the bank provides liquidity and transaction services to the sellers;

(v) these bank's services result in risk exposure for the bank in terms of both solvency and liquidity risks, both of which the bank must manage;

(vi) the bank's services reduce the buyers' transaction costs, and mitigate risk exposure and other transaction costs to sellers; and

(vii) as a result of the services and guarantees that the bank provides, the bank's deposit liabilities function as a generally accepted medium-of-exchange cash substitute.

This description of banks as monetary intermediaries has at least five implications for the role of banks in the monetary system.

The first implication is that bank deposits (bank-money) are not created out of nothing. Rather they are the result of value-for-value transactions between economic agents transacting on the basis of debt-credit contracts. The second is that the rationale for the existence of a bank is the reduction of the various risks and costs associated with the intertemporal exchange of G&S. The third is that the quality of bank-money is underwritten by a portfolio of private debt obligations. A bank's lending decisions therefore impact on the quality of the money supply, which therefore has the potential to influence broader economic and financial stability. The fourth is that the growth of bank deposits is expected to be positively correlated with general economic activity. This provides an explanation for the empirical observation that bank intermediation and GDP have strong positive correlations. Finally, bank loans are not dependent on prior savings. Rather, they depend on factors that include: the liquidity needs of economic agents and their desire to hold bank deposit liabilities as a medium-of-exchange; the bank's administrative expertise and ability to manage and reduce the risks and costs related to loan default and deposit withdrawal; and the bank's expectations regarding its profitability and solvency in extending credit to economic agents. These last two points suggest that bank credit extension is likely to be strongly pro-cyclical. This suggests that macro-prudential policy has an important role to play in regulating financial cycles. In Part 2 of this chapter, I model the role of banks as monetary intermediaries.

Part 2: Modelling the Role of Banks as Monetary Intermediaries

2.9. Introduction

Bank demand-deposit liabilities are fungible cash substitutes that are convertible to cash on-demand-and-at-par. But the value of these liabilities is backed (for the most part) by asset portfolios consisting of risky, illiquid, private debt. This balance sheet structure makes banks vulnerable to “runs”, and leaves the financial system vulnerable to crises. Why then do banks perform this function with this balance sheet structure if it means they are fragile? This is an enduring puzzle within bank theory.[23] I propose a new answer to this puzzle. I argue that bank balance sheet structure reflects the uniquely monetary function of banks: the transformation of debt into money in order to mitigate various risks and other transaction costs that impede the intertemporal exchange of G&S between buyers and sellers. This structure is prone to fragility when provisions for the wide array of shocks that arise from the broader (uncertain) institutional operating environment that can affect this function, are inadequate. This answer is useful because it is not limited to asymmetric information and moral hazard problems. Instead, it accounts for bank failure in the more general terms of the broader theory of the firm. This theory is additionally compelling because it provides solutions to a number of other puzzles within bank theory literature.

[23] As discussed in Gu, et al. (2019).

In Part 2 of this chapter, I extend the discussion from Part 1 and develop a model of intertemporal trade between buyers and sellers. I use this model to derive the essential characteristics of bank balance sheets from first principles. The model illustrates that bank balance sheet structure and characteristics are a direct consequence of two phenomena. The first is the role of banks in facilitating transactions between buyers and sellers by transforming risky debt-credit contracts into money. The second phenomenon is the broader institutional environment in which banks operate. I show that monetary intermediation is both more efficient, and more effective, in institutional environments where markets are *more* complete, and where information asymmetries and other market imperfections are *reduced*. I show that, as monetary intermediaries, banks are more fragile the more *imperfect* the broader institutional environment, *ceteris paribus*. In terms of this framing, bank runs are a failure of banks to adequately account for, and respond to, shocks characterising the broader institutional environment. In this regard, banks are not *designed* to be fragile,[24] but like any other entity that must operate and respond to changes in its environment, a failure to do so adequately and appropriately may lead to their demise.[25]

This second part of the chapter is structured as follows: Section 2.10 provides an overview of the model; Section 2.11 develops the model; Section 2.12 discusses the implications of the model; and Section 2.13 concludes.

[24] Cf. Calomiris & Kahn (1991) and Flannery (1994).

[25] The question for policy makers in this case is how to address this market failure of inadequate provisioning. This is the ambit of prudential policy.

2.10. Model overview

The model is a simple three-period trade model. There are buying agents and selling agents. At date 0, sellers receive an endowment of consumable G&S, but only wish to consume at date 2. Buyers receive no consumable G&S at date 0, but wish to consume at date 0. Buyers receive an endowment of consumable G&S at date 2, and can credibly commit to deliver their date 2 endowment of G&S in exchange for the date 0 G&S endowments of sellers. Market infrastructures exist, and therefore buyers and sellers can enter into market exchange contracts structured as follows: at date 0, buyers purchase G&S from sellers by entering into debt-credit agreements with sellers; in terms of these contracts, buyers agree to provide sellers with G&S at date 2 in return for the G&S received from sellers at date 0. G&S are therefore intertemporally exchanged by means of debt-credit agreements. Sections 2.11.1 through to 2.11.3 present this baseline setup.

Following from the baseline setup, an alternative medium-of-exchange is introduced in Section 2.11.4 in the form of cash. Two potential problems with this setup are then considered. First, the availability of cash may be constrained relative to the efficient volume of trade given the distribution of resources and agent preferences. Second, while debt-credit agreements still provide an alternative medium-of-exchange to cash, buyers' date 2 G&S endowments may be subject to shocks originating from the broader institutional environment in between date 0 and date 2. These shocks may mean that buyers are unable to deliver the agreed upon quantities of G&S at date 2 in terms of the debt-credit agreements. These shocks are characterised as "bad luck" as opposed to "bad faith". This risk undermines the efficiency of debt-credit agreements as media-of-exchange. I explore the characteristics of mechanisms that endogenously resolve these two problems and thereby increase the efficiency of the exchange process. These mechanisms mitigate the risks and costs (relative to the constant value of cash) of using debt-credit agreements as media-of-exchange. They do so by transforming the debt-credit agreements into riskless fungible cash substitutes. I refer to these mechanisms as *monetary mechanisms*. The reduction of risks and costs by these mechanisms increases the efficiency of trade.

The discussion considers three different scenarios characterised by different *shock types* and by different underlying institutional *environments*. In the first scenario (Section 2.11.5), I consider an expected shock with a known risk of repayment variance. Agents transact in an environment where markets are complete and competitive; transacting is otherwise frictionless; information is symmetric; and agents are risk averse. A mechanism that (i) pools and diversifies the risk of the debt-credit contracts, and (ii) produces riskless derivative financial instruments that are underwritten by these debt-credit contracts, improves the Pareto efficiency of exchange.[26] I show that these derivative financial instruments will trade at par with cash and will function as a first-best cash substitute.

In the second scenario (Section 2.11.6), I consider expected shocks that result in repayment variance, some of which can be simply mitigated through pooled diversification, and the remainder of which is insurable through complete insurance markets. Once again, a mechanism that provides derivative financial instruments underwritten by the debt-credit contracts and insured against these shocks results in

[26] This mechanism could be a market but does not need to be.

Pareto efficient exchange. This mechanism passes the undiversified risk of the shock on to third parties, and the derivative instruments trade at par with cash. This monetary mechanism has many important bank-like features.

In the third scenario (Section 2.11.7), I change the characteristics of the broader institutional environment to include incomplete or otherwise imperfect markets and market transaction costs. In this environment, the characteristics of efficient monetary mechanisms, which perform the monetary function of transforming debt-credit agreements into derivative financial instruments that trade on par with cash, have the features of what in practice constitute banks. I illustrate that the specific balance sheet composition of these monetary mechanisms depends on the characteristics of the broader institutional environment. This finding is in accordance with the *functional perspective* approach of financial intermediation theory.[27] The functional perspective argues that institutional form follows from economic function. From this perspective, if the *environment* changes, the balance sheet structure of a monetary intermediary must adjust accordingly to perform its intended function. Like any other entity existing within a broader environment, a monetary intermediary is therefore vulnerable to changes in the broader environment if it is unable to adjust appropriately in response to these external changes. Accordingly, I show that bank resilience is promoted through measures that protect the solvency of a bank's balance sheet from external environmental shocks that negatively affect the value of its asset portfolio. These measures include provisions for negative impacts through either on-balance-sheet items, such as higher equity capital, or off-balance-sheet items, such as collateral and

[27] See Merton (1993; 1995) and Merton & Bodie (1995; 2005).

securitisation. Additional measures can include regulation aimed at minimising bank balance sheet exposure to undiversified and pro-cyclicality risk; restricting banks to high-quality asset portfolios; and regulation aimed at encouraging development towards more complete and informationally perfect markets (Section 2.11.8). The model is useful for explaining the specificity of bank balance sheets, products, and contracts. It also provides new and compelling answers to a number of puzzles within bank theory.

2.11. The model

In this section, I develop a model of monetary intermediation. It is a three-period partial equilibrium endowment model. It is centred on the exchange of G&S across time using debt-credit financial instruments as media-of-exchange. There are two core design features of the model. The first core design feature is the temporal wedge separating the exchange of G&S. The second is that this time wedge introduces risks and uncertainties. The model derives the monetary intermediation role of banks as an endogenous mechanism for mitigating the costs of these risks and uncertainties.

2.11.1. Setup: Agents, endowments, preferences, timing, technology

The initial setup of the model is as follows.

(i) There are three discrete time periods represented by numbered subscripts $t_{\tau \in \{0, 1, 2\}}$.

(ii) There are a total of N agents consisting of two preference types, $j \in \{b, s\}$. Type b agents are "buyers" of current G&S, and type s agents are "sellers" of current G&S. Agents are aware of their types and this information is symmetric.[28] N is a statistically large number.

(iii) Some proportion $0 < \psi < 1$ of agents are buyers, and the remaining $1 - \psi$ are sellers.

(iv) There are three *perfectly divisible* and tradable property types, $i \in \{C, K, M\}$: consumable G&S (C); debt-credit contracts obliging delivery of C in t_2 (K); and cash (M), a financial instrument that converts into C on a one-to-one basis in t_2.[29]

(v) Agents each receive endowments of property (e_{ji}). Sellers each receive a unit value endowment of C in t_0 ($e_{sC} = 1$). Buyers each receive an endowment of ρ units of C in t_2 and can enter into a unit value of debt-credit contracts in t_0 ($e_{bK} = 1$). A unit value of K can therefore return a maximum of $\rho_{K(C)}$ units of C in t_2.[30] Initially ρ has certain value—this restriction is later relaxed.[31]

(vi) Initially, there are no endowments of M (i.e., $e_{jM} = 0$)—this restriction is later relaxed.

(vii) Buyers initially receive utility (U) in t_0 by consuming the quantity of C purchased from sellers (Q_C) such that $U_b = f(Q_C) = Q_C$. Sellers initially

[28]Unlike a common setup within financial intermediation literature where type is revealed as a shock, agents in this model are aware of their type from the outset and do not require insurance against this shock. Unlike a typical mechanism design setup, the problem of agents concealing their type is not of concern here either.

[29] Cash in the model is an exogenous, publicly issued financial instrument with the standard properties of cash as described in, for example, Knapp (1924), Gurley & Shaw (1960), or Goodhart (1998).

[30] Subscripts on price or return variables are used to denote measurement in "units of"; e.g., $P_{C(M)}$ is read "the price of G&S in units of cash"; $\rho_{K(C)}$ is read "the return on private debt-credit contracts in units of G&S"; etc.

[31] ρ is the future payoff associated with K. The value of ρ is taken as exogenous in this model. In the DSGE model in Chapter 4, ρ is endogenous.

receive utility in t_2 from the quantity of debt-credit contracts entered into in t_0 such that $U_s = g(\rho Q_K) = \rho Q_K$. Initially, endowments of t_0 C can only be consumed in t_0 and cannot be carried over to t_1 or t_2—this restriction is later relaxed.

In addition to this setup, price stability (as measured by the purchasing power of cash) is assumed to be a desirable feature of the economy.[32] Information is symmetric.[33] The setup applies to the immediate discussion, but as the discussion progresses, the following assumptions are relaxed: the assumption preventing intertemporal substitution (termed Assumption 1 in what follows);[34] the absence of cash (Assumption 2); and the certainty of ρ (Assumption 3). When relaxing Assumption 3, ρ is subject to an information shock in period t_1. Figure 2-3 (i) and (ii) summarises the sequence of events in the model. In Section 2.11.2, I establish the baseline setup of intertemporal exchange using debt-credit contracts.[35]

[32] This assumption is justified on the basis of theory. Wicksell (1936 [1898], p. 4) writes: "It follows that if it were in our power to regulate completely the price system of the future, the ideal position … would undoubtedly be one in which, without interfering with the inevitable variations in the relative prices of commodities, the general average level of money prices … would be perfectly invariable and stable". Woodford (2003, pp. 2-3) writes: "The recognition that the purchasing power of money need not be dictated by any 'natural' market forces, and is instead a proper subject of government regulation, … should in principle make possible greater stability of the standard of value, facilitating contracting and market exchange … [because] price stability creates a sound basis for real economic performance". As empirical evidence of this desirability, Brunner & Meltzer (1971, p. 803) state: "Our analysis suggests an extension of Gresham's law … to the case of multiple mediums of exchange …, stable monies drive out variable monies". See Borio (2019) for further discussion on this point.

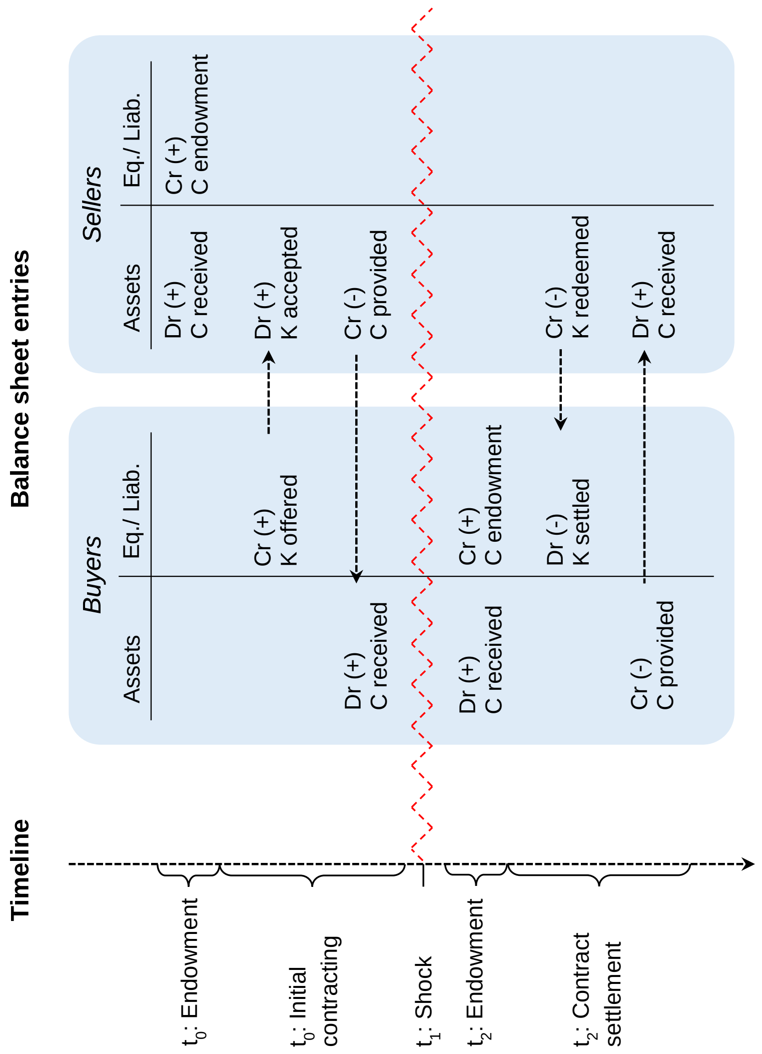
[33] While information is symmetric, some information may be unknown. The fundamental information friction is therefore not asymmetry of information, but rather *uncertainty* of information.

[34] When relaxing Assumption 1, G&S are no longer restricted from carrying over from t_0 to t_2, and agents are no longer limited to deriving linear utility from only a single period's consumption.

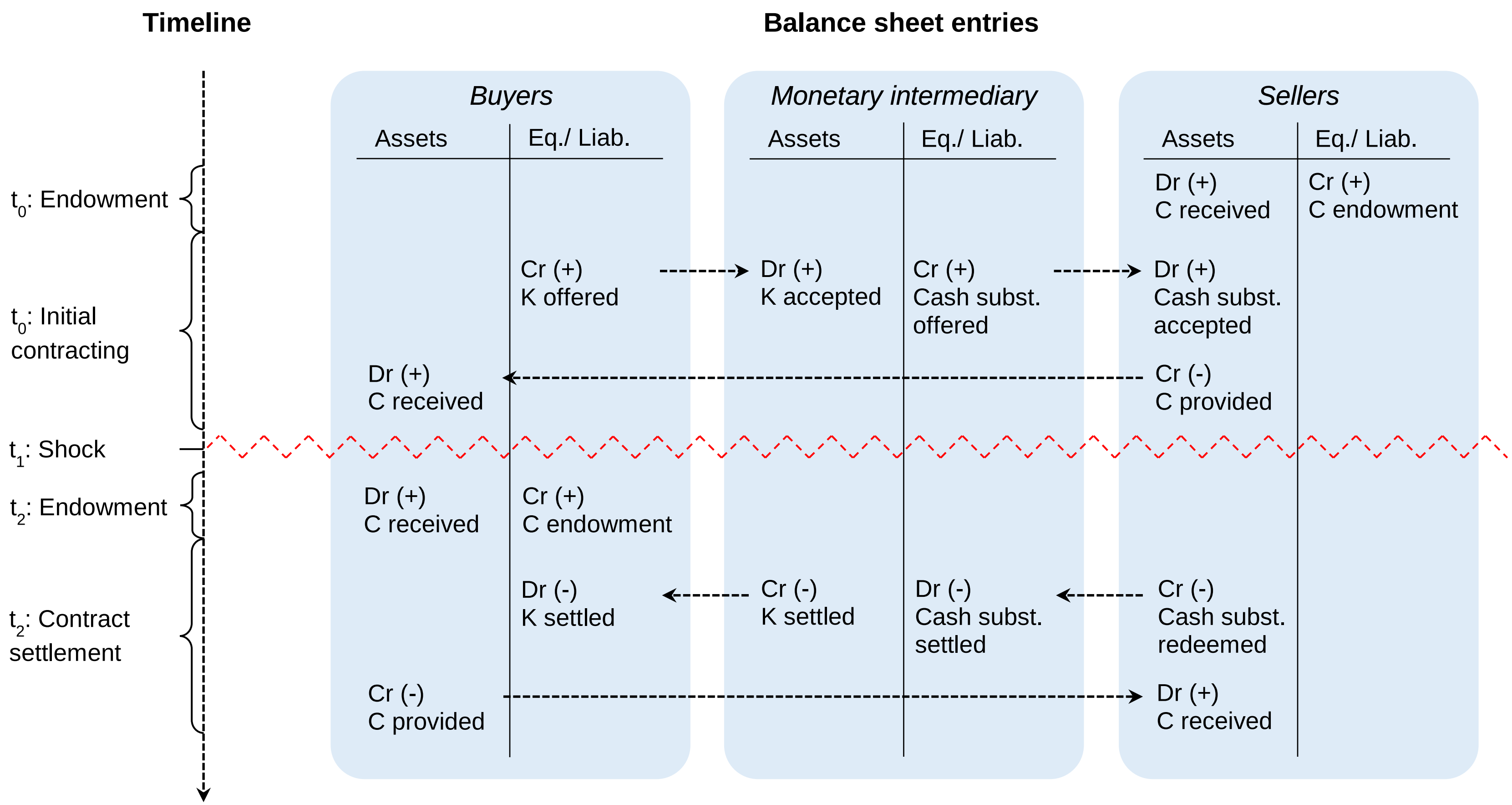
[35] The baseline setup reflects a pre-monetary exchange environment where debt-credit contracts are used as media-of-exchange—see for example von Reden (2007). The discussion in Section 2.2 motivates this setup.

Figure 2-3: Model sequence of events and agent balance sheets over time

(i) Sequence of events and balance sheet composition over time *without* monetary intermediation



(ii) Sequence of events and balance sheet composition over time *with* monetary intermediation



2.11.2. The baseline model of exchange

In t_0, a perfectly competitive Walrasian market exchange opens and agents are able to trade their endowments with one another. K is both offered and accepted in exchange for C in t_0. The equilibrium price of a unit of K in terms of units of C $\left(P_{K(C)}\right)$ is:[36]

$$P_{K(C)} = \frac{1-\psi}{\psi} \equiv \overline{P}_{K(C)}. \tag{2.1}$$

At equilibrium, total endowments of both C and K are traded. I denote an equilibrium price with a bar ($\overline{P}$) if it has the characteristic that *total* endowments are traded—this notion is important in Section 2.11.4. The outcome of this process is that non-utility-yielding G&S are exchanged (by means of K) for utility-yielding G&S. K serve as media-of-exchange. It is straightforward to demonstrate that both individual utility and social welfare are maximised through trade.[37] This is a clear and intuitive incentive for avoiding autarky. Equilibrium allocations are Pareto optimal. Agents self-select their roles according to their preferences, and exchange is incentive compatible. With no information asymmetries and no trading restrictions, exchange takes place with no need for intermediation, and final allocations are not improved by coalitions or by prior contracting outside of this exchange environment. There are no intertemporal trade-offs, meaning that the price of K is independent of ρ.[38]

---

[36] See Appendix 2B for the derivation.
[37] See Appendix 2B.
[38] From Equation (2.1), the equilibrium quantity of C exchanged per unit of K is simply the proportion of sellers to buyers. Therefore, in this setup, ρ is a trivial value of consumption in t_2.

<u>LEMMA 1</u>: In the baseline model, there are no information asymmetries, no trading restrictions, and no intertemporal trade-offs. The intertemporal exchange of G&S between utility-maximising buyers and sellers takes place by means of simple bilateral debt-credit financial contracts. These contracts represent claims against future G&S. The intertemporal equilibrium is Pareto optimal.[39]

2.11.3. Relaxing Assumption 1: Introducing intertemporal substitution

Two changes to the baseline setup now occur. First, agents are now assumed to have standard and symmetric consumption smoothing preferences. Second, C can be carried across periods; this gives sellers the choice of selling their endowment of t_0 C, or costlessly carrying their endowment forward and deriving utility by consuming their endowment in t_2. With intertemporal substitution, unconstrained rational agents will trade at prices that equate their intertemporal marginal rates of substitution.[40] The incentive compatible, Pareto efficient, equilibrium price of K in a perfectly competitive market is therefore determined by ρ such that:

$$P_{K(C)} = \rho_{K(C)}. \tag{2.2}$$

Equation (2.2) represents a standard financial asset pricing result.[41]

[39] Baltensperger (1989, pp. 97-98) explains the intuition behind this lemma, writing: "The credit market is essentially a market for intertemporal exchange. Something is given up in the present in exchange for something else in the future—or *vice versa* … The possibility of intertemporal exchanges allows market participants the realization of utility gains, just as voluntary exchange in general is mutually advantageous. The basic reason for this is that individuals are not normally indifferent about the distribution or their consumption over time but care about it".

[40] For parsimony, utility is assumed to be additively separable over consumption streams, and subjective rates of discount and depreciation are both zero. These elements are easily changed, but add needlessly to the complexity of the model with little added insight.

[41] The discussion in Jacklin (1987, p. 34) provides a detailed discussion of this pricing result.

<u>LEMMA 2</u>: With intertemporal substitution and standard consumption smoothing preferences, sellers sell G&S in the current period iff the future utility provided by the financial instrument received in payment for the G&S sold yields at least as much future utility as the future utility from carrying the G&S over to future periods.

If there is some transaction cost (ε) associated with using *K* as media-of-exchange, then Equation (2.2) can be adjusted to take account of this such that:

$$P_{K(C)} + \varepsilon = \rho_{K(C)} \Leftrightarrow P_{K(C)} = \rho_{K(C)} - \varepsilon. \tag{2.3}$$

Equation (2.3) formalises the result that market frictions and transaction costs lower the amount of *C* per unit value of *K*. Consistent with the transactions theory of the firm (Coase, 1937), the welfare effects of transaction costs and other frictions explain the existence of mechanisms able to *reduce* these costs. Mechanisms that can reduce these costs therefore mitigate the reduction in value of *K*, improve utility for buyers, and achieve a Pareto improvement.

<u>LEMMA 3</u>: Market frictions and other transaction costs characterising the use of debt-credit contracts as a medium-of-exchange lower the marginal return on these contracts. A mechanism that reduces these costs and frictions mitigates the reduction in price of these contracts. Such a mechanism is both Pareto improving and incentive compatible.

2.11.4. Relaxing Assumption 2: Introducing cash [42]

Transactions have not yet involved cash, and bilateral private debt-credit contracts have served directly as the media-of-exchange. Endowments of cash (M) are now introduced. In t_0, each seller receives an endowment $e_{sM} = \alpha$.[43] A unit of M converts to a unit of C in t_2 with certainty (i.e., prices are assumed to be constant over time).[44] M can be traded for K (in complete and frictionless financial markets), or C (in complete and frictionless goods markets). To eliminate arbitrage across goods and financial markets, the ratio of exchange of M for K will be:

$$1M : \rho^{-1}K \Leftrightarrow \rho M : 1K.$$

At this price ratio, ρ units of M provides a perfect medium-of-exchange substitute for a unit value of K. This result is a standard asset pricing result for cash in the general equilibrium literature.[45] At this exchange rate, sellers are indifferent between receiving either one unit of M or ρ^{-1} units of K in payment for one unit of C.

[42] The focus in this thesis is not on *why state issued cash* is the *de facto* medium-of-exchange. Both Knapp's (1924) "Chartalism", and Wray's (2012; 2014) "Modern Money Theory", offer explanations that are compatible with this model for why state issued cash is the *de facto* medium-of-exchange. The model assumes that an instrument with the properties of cash *is* the generally accepted medium-of-exchange. A discussion on why the properties of cash are important as a medium-of-exchange is beyond the scope of the analysis, but Brunner & Meltzer (1971) provide a relevant discussion. An underlying assumption of the model is that the characteristics of cash make it a useful medium-of-exchange.

[43] Sellers (as opposed to buyers) are endowed with cash so that buyers must go through financial markets to obtain M. This provides a role for financial markets in the model.

[44] The quantity α assumes a constant velocity of circulation for ease of calculations. By fixing M across periods, the real value of money across periods is fixed and monetary prices are therefore stable—that is, there is no inflation in the model. Extending the model to include inflation is straightforward. Inflation will reflect as a variable nominal return on K or variable real return on M. Section 2.11.6 considers a variable nominal return on K, and so the results of Section 2.11.6 offer a first approximation of the effects of nominal instability.

[45] See Walsh (2010). This result is also referred to as the cost of liquidity services (Svensson, 1985).

The quantity of cash (α) relative to the Pareto optimal level of intertemporal exchange matters for determining the media-of-exchange used to trade. The quantity of cash is *scarce* relative to the Pareto optimal level of exchange if $0 < \alpha \ll 1$ such that:

$$P_{K(C)} = P_{K(M)} = \rho_{K(M)} > \overline{P}_{K(M)} = \alpha \overline{P}_{K(C)}. \tag{2.4}$$

Equation (2.4) is important.[46] Equation (2.4) specifies circumstances where using only M as the medium-of-exchange is not Pareto optimal given the quantity of K available (and keeping the purchasing power of cash constant across time). A unit value of K will be priced to be perfectly substitutable with a unit value of M and provides a medium-of-exchange substitute for M. In this case, using private debt-credit contracts as a medium-of-exchange alongside cash is Pareto improving.[47]. Buyers are therefore not liquidity constrained in any traditional sense by the scarcity of cash.

DEFINITION 1: A cash substitute *is a financial instrument with the same risk characteristics and intertemporal purchasing properties as cash.*

[46] See Appendix 2C for the derivation and further discussion. The circumstance of limited cash is of interest because casual empiricism indicates that cash forms only a small portion of money in circulation within modern economies—only 3% in the UK (McLeay, et al., 2014b). But Equation (2.4) can also be interpreted as a circumstance where converting debt directly into cash is costly, and 1- α is the value of these costs. In this case, using debt-credit contracts as a direct cash substitute circumvents these costs. Shubik (2016, p. 11) writes: "As soon as we observe that the execution of a trade involves some finite amount of time the cash flow conditions involving the use of a money become relevant. If these conditions hamper trade an economy devises institutions and instruments to cure this. Even at its simplest there are many ways to do so. One might avoid the need to create credit by speeding up the velocity of payment, but this, while not impossible, is difficult to do and both habit and technology impose an upper bound on velocity. Most cures involve the invention of near monies or credit". Similarly, Wallace (1990), Azariadis, et al. (2001), and Bullard & Smith (2003a; 2003b) show that privately issued circulating liabilities improve resource allocations in markets where cash availability is restricted or costly.

[47] There is no need for an intermediary in this setup. But as per Lemma 3, the introduction of some costly friction, and the minimising of this friction through intermediation, provides motivation for the existence of an intermediary.

LEMMA 4: If the quantity of cash is insufficient for exchanging the Pareto efficient quantity of G&S, then using debt-credit contracts as a cash substitute improves the efficiency of exchange.

2.11.5. Scenario 1—Relaxing Assumption 3: Return variance and risk sharing

The discussion thus far has assumed that buyers' future endowments of C (i.e., ρ) are known with certainty. Assume now instead that ρ is subject to a shock in period t_1, and there are two potential return outcomes: a high outcome (ρ_H) and a low outcome (ρ_L). Assume individual endowments of K earn a distribution with known i.i.d., yielding either ρ_H with probability θ or ρ_L with probability $(1-\theta)$. Assume agents are *symmetrically unaware* of the actual outcome any particular endowment of K will yield and hence there is a degree of heterogeneity pertaining to these otherwise indistinguishable financial instruments. The risk concern of sellers is therefore not that a particular unit of K will return ρ_L due to asymmetric information (e.g., bad faith),[48] but rather that some misfortune will prevent ρ_H (e.g., bad luck).[49] There are therefore

[48] Cf. Diamond (1984) and Boyd & Prescott (1986).
[49] Agents are in effect assumed to be acting in good faith; agents do not know *a priori* if they hold good capital (high return) or bad (low return). Instead, low returns can be thought of as a misfortune (unanticipated loss of income or poor sales revenue are concrete examples of this), but no agent has reason to believe in the current period that they are more or less likely to fall victim to this misfortune. I assume here that the broader institutional environment has in place mechanisms sufficient for discouraging generally fraudulent behaviour. This benign framing of the default risk faced by banks has its roots in the discussion of Thornton (1965 [1802]), who notes that the fundamental starting point of trade in general relies on both the legal institutional infrastructure and the good faith of the parties involved. Thornton (1965 [1802], pp. 75-76) writes: "Commercial credit may be defined to be that confidence which subsists among commercial men in respect to their mercantile affairs … Even in that early and rude state of society, in which neither bills nor money are as yet known, it may be assumed, that if there be commerce, a certain degree of commercial credit will also subsist … In a society in which law and the sense of moral duty are weak, and property is consequently insecure, there will, of course, be little confidence or credit, and there will also be little commerce. This commercial credit is the

neither agency problems nor informational asymmetries in this setup. This is a departure from the asymmetric information setup that is common to financial intermediation literature. The expected yield $\mathrm{E}(\rho)$ and known variance σ^2 of any given unit of K is therefore given as:

$$\mathrm{E}(\rho) = \theta\rho_H + (1-\theta)\rho_L$$

$$\sigma^2 = \theta\big(\rho_H - \mathrm{E}(\rho)\big)^2 + (1-\theta)\big(\rho_L - \mathrm{E}(\rho)\big)^2.$$

Equation (2.2) can be re-specified as:

$$P_{K(C)} = \mathrm{E}\big(\rho_{K(C)}\big). \tag{2.5}$$

If agents are risk averse, and letting ε' be the risk premium required for a certainty equivalent level of cash utility,[50] Equation (2.5) can be written:

$$P_{K(C)} = P'_{K(C)} + \varepsilon' = E\big(\rho_{K(C)}\big) \Leftrightarrow P'_{K(C)} = E\big(\rho_{K(C)}\big) - \varepsilon' \tag{2.6}$$

where $P'_{K(C)}$ is the *certainty equivalent price* of K, and reducing exposure to risk premium ε' increases $P'_{K(C)}$ towards the risk-free price $P_{K(C)}$ with $P_{K(C)} > P'_{K(C)}\ \forall\ \varepsilon' > 0$. From Equations (2.4) and (2.6), using K and M as substitutable media-of-exchange is Pareto improving if:

$$P_{K(M)} > P'_{K(M)} = E\big(\rho_{K(M)}\big) - \varepsilon' > \overline{P}_{K(M)}. \tag{2.7}$$

Sellers are indifferent between the lower $P'_{K(M)}$ for a unit of undiversified K and higher $P_{K(M)}$ for a diversified and riskless derivative unit of K. However, the assumption of

foundation of paper credit; paper serving to express that confidence which is in the mind, and to reduce to writing those engagements to pay, which might otherwise be merely verbal."

[50] Here, certainty equivalent utility has the standard meaning. It is the level of utility associated with a payoff that is certain and not subject to any risk.

complete and frictionless competitive exchange environments ensures Pareto efficiency by driving $P'_{K(M)}$ towards the risk-free price $P_{K(M)}$ through pooled diversification.[51] Pooled diversification can be achieved using various pooling strategies either with or without the use of intermediaries.[52] Equation (2.7) is a monetary mechanism.

<u>DEFINITION 2</u>: *A* monetary mechanism: (i) *pools debt, and (ii) issues derivative financial claims against the underlying portfolio of pooled debt. These claims are readily transferable, homogenous and risk free (relative to cash). They provide a fungible cash substitute.*

2.11.6. Scenario 2—Relaxing Assumption 3: Systematic risk

If the return on a diversified portfolio of K is subject to systematic risk relative to cash, and pooling is unable to eliminate this risk—for example a shock that lowers $\mathrm{E}(\rho_{K(C)})$[53]—then following from Equation (2.7), a derivative unit of diversified debt (that is exposed to this systematic risk) will trade at a discount relative to the *riskless* cash price $(P_{K(M)})$, such that:

$$P_{K(M)} > P''_{K(M)} = \mathrm{E}(\rho_{K(M)}) - \varepsilon'' > \overline{P}_{K(M)} \quad (2.8)$$

---

[51] Although the quantity of M is limited compared to the Pareto efficient quantity of exchange, with no uncertainty, no transaction costs, and no-arbitrage opportunities, there is no advantage to holding M as opposed to holding a claim against a diversified portfolio of K. Under these circumstances, holding a claim on the insured diversified portfolio has the properties of a cash substitute. Agents are therefore indifferent between holding M and holding a claim against the portfolio.

[52] See Appendix 2D for further discussion on these strategies.

[53] A systematic event, for example, could reduce either ρ_H, ρ_L, or θ and so reduce $\mathrm{E}(\rho_{K(C)})$.

where $P''_{K(M)}$ is the *certainty equivalent* price of *diversified* private debt; ε'' is the cost of insuring against systematic risk in complete and frictionless financial markets; and $P_{K(M)} = P''_{K(M)} + \varepsilon'' = P'_{K(M)} + \varepsilon' + \varepsilon'' = \mathrm{E}(\rho_{K(M)})$.

Equation (2.8) is a monetary mechanism that takes into account a broader universe of costs and risks than Equation (2.7). To clarify what Equation (2.8) means, it can be translated back to the example of a transaction between Ms S and Mr B in Part 1 of this chapter as follows. The agreed purchase price for Ms S' property was $100,000. To purchase the property, Mr B's debt must be transformed into a derivative financial instrument with a guaranteed money value of $100,000 $(P''_{K(M)})$. If the cost of guaranteeing the value of Mr B's debt is 5% of the guaranteed value, or $5,000, then Mr B must enter into a debt-credit contract with an expected value of $105,000 $(\mathrm{E}(\rho_{K(M)}) = \$100{,}000 + \$5{,}000)$. The contract value of $105,000 includes the costs of guaranteeing Mr B's debt to the value of $100,000. In a competitive environment, the diversifiable risk (ε') will be diversified away, lowering the price of the contract B must enter into in order to fund the purchase the property. If the undiversifiable cost of the guarantee (ε'') is, say, only 1% of the value, then Mr B would only need to enter into a contract with an expected value of $101,000. Competition is expected to drive this pricing process.

2.11.7. *Scenario 3—Cash-on-demand-and-at-par, uncertainty, and moral hazard*

Taking Equation (2.8) as a starting point, bank balance sheet structure and contracts can be interpreted in terms of the interactions of a monetary mechanism with the broader institutional environment. The following discussion illustrates this point. Consider, for example, that it is incentive compatible for *buyers* that the liabilities of a monetary mechanism be convertible to cash on-demand-and-at-par. As Brunner & Meltzer (1971, p. 803) write:

> "[G]overnment or private producers desirous of maintaining the circulation of government and privately produced monies have found it desirable to provide arrangements for exchange of one money for the other on demand. More importantly, issuers of privately produced money expand the market for their product by maintaining a fixed exchange rate. By doing so, they lower the users' cost of information and increase the demand for their product."

In other words, conversion on demand at a fixed exchange rate expands the market for the liabilities of the monetary mechanism, thereby increasing the price buyers receive for their debt, which forms the asset side of the balance sheet of the monetary mechanism.

Consider further, that the option to convert claims against a diversified asset portfolio into cash on-demand-and-at-par is also incentive compatible from the perspective of *sellers*. A cash on-demand-and-at-par option serves as a put option on the underlying portfolio (Merton, 1977), and this provides insurance against ε' and ε'' to sellers. Further, a cash-on-demand-and-at-par guarantee insures sellers against *any* risks, costs, frictions, and other uncertainties that would otherwise reduce the value of

K relative to M (henceforth these factors collectively are ε''').[54] Equation (2.8) can be adjusted to take account of these factors such that:

$$P_{K(M)} > P'''_{K(M)} = E\big(\rho_{K(M)}\big) - \varepsilon''' > \overline{P}_{K(M)} \tag{2.9}$$

where $P_{K(M)} = P'''_{K(M)} + \varepsilon''' = \mathrm{E}\big(\rho_{K(M)}\big)$.[55] Equation (2.9) is a monetary mechanism that is incentive compatible and Pareto improving if it: (i) insures sellers against ε'''; and (ii) minimises the costs ε''' incurred by buyers. With regards to (i), a cash-on-demand-and-at-par guarantee insures sellers against ε'''. With regards to (ii), methods for minimising ε''' incurred by buyers will depend on the nature of ε''' and the broader institutional environment. Where insurance markets are incomplete, or the risks of ε''' are otherwise uninsurable through markets, costs borne by buyers or sellers can be minimised through the internal balance sheet features of the monetary mechanism (such as equity capital),[56] or through broader institutional arrangements (such as regulation).[57]

---

54 These factors include risks, costs, and frictions associated with incomplete markets, asymmetric information, and time discount factors, such as the need for liquidity as defined in Diamond & Dybvig (1983), and problems with commitment, for example, Kocherlakota (1998) and Calomiris & Kahn (1991).

55 Equation (2.9) is analogous to discounting a bill of exchange or other promissory notes as described by Thornton (1965 [1802], pp. 82-89), Macleod (1866), and Wicksell (1936 [1898]). Using the terminology relevant to this early period, one could say that a bill for an amount of $\mathrm{E}\big(\rho_{K(M)}\big)$ is discounted by some amount ε''' and exchanged for bank notes to the value of $P'''_{K(M)}$.

56 This is consistent with Knight (1964 [1921], pp. 46-47) who writes: "The fact is that while a single situation involving a known risk may be regarded as 'uncertain', this uncertainty is easily converted into effective certainty; for in a considerable number of such cases the results become predictable in accordance with the laws of chance, and the error in such prediction approaches zero as the number of cases is increased. Hence it is simply a matter of an elementary development of business organization to combine a sufficient number of cases to reduce the uncertainty to any desired limits. This is, of course, what is accomplished by the institution of insurance … If it is objected that practical difficulties may prevent insurance even where the risk is determinate, the reply is that insurance, in the technical sense, is only one method of applying the same principle. We shall show at length in our general discussion of risk and uncertainty that if the risk is measurable, but the 'moral factor' or some other consideration makes ordinary insurance inapplicable, some other method of securing the same result will be developed and employed".

57 See for example Fama (1980, p. 42), DeAngelo & Stulz (2015), Brunnermeier & Sannikov (2016), Phelan (2016), and Merton & Thakor (2018) for discussions on the rationale for deposit insurance.

If moral hazard on the part of buyers is of concern, this can be explicitly curtailed by building individual rationality constraints into the monetary mechanism, such as in Kehoe & Levine (1993), or by introducing collateral requirements, such as in Holmström (2015).[58] Additionally, ε''' may be mitigated through technology improvements for more effectively screening and monitoring assets, or through derivative products that efficiently shift risks off balance sheet (such as securitisation). These features tend to be characteristic of bank balance sheet structures, products, and contracts in practise.[59]

<u>PROPOSITION 1</u>: Where G&S are exchanged on the basis of idiosyncratically risky private debt, a monetary mechanism: (i) mitigates risk; (ii) provides debt issuers with the highest incentive compatible price for their private debt; and (iii) maximises social welfare by: (I) pooling the debt to form a diversified portfolio; (II) insuring the portfolio against systematic risk; (III) issuing homogenous derivative financial instruments that represent riskless (relative to cash) claims against the portfolio; and (IV) providing infrastructure that enables these derivative financial instruments to be transferred directly as a medium-of-exchange cash substitute.

[58] A perfectly competitive market has been assumed up to this point. This assumption necessarily implies assigned property rights and the presence of legal and institutional structures designed to enforce these rights, and to curtail problems of moral hazard or bad faith (Goodhart, 1998). This institutional infrastructure is presumably a reasonable assumption given that anthropological evidence indicates these sorts of institutional structures were prerequisites of exchange relationships (Greif, 1993). Moreover, Gorton & Winton (2003) find a "lack of compelling evidence" that moral hazard is a problem in banking contracts. And indeed, both Thornton (1965 [1802]) and Wicksell (1936 [1898]) indicate that modern banking arose from relationships based on *trust*. The tradition of framing the *raison d'être* of banks as a means of endogenising constraints against bad faith therefore misses this context, and takes an extreme viewpoint by ignoring the constraints imposed by the broader institutional environment. By contrast, the setup used here has assumed that appropriate exogenous institutional structures govern agency relations. This assumption could been seen as taking an equal but opposite extreme, in which case, extending the model to endogenise constraints against issues of moral hazard would address this concern.

[59] If the costs of borrowing cash directly from other economic agents (who have cash available to lend) is more than the cost of ε''' incurred by buyers using a monetary mechanism to convert debt into a cash substitute, then using the monetary mechanism will be preferable to borrowing cash directly.

2.11.8. *Banks, runs, and crises*

From a *functional perspective*, institutional form follows from function (Merton, 1993; 1995; Merton & Bodie, 1995; 2005). From this functional perspective, a bank can be explained as the institutional form of the monetary mechanism described by Equation (2.9). From this perspective, a bank is the institutional form of a monetary mechanism that performs the function of transforming buyer debt into a cash substitute for sellers. From Equation (2.9), the objective function of a bank as monetary intermediary can be specified as:

$$\min \mathrm{E}\big(\rho_{K(M)}\big) = P'''_{K(M)} + \varepsilon''' \tag{2.10}$$

where the value of $P'''_{K(M)}$ (i.e. the deposit value) is fixed and guaranteed.[60] In other words, the bank's objective function is to minimise costs ε''' thereby minimising the amount of debt buyers must pledge for a guaranteed deposit value. By minimising ε''', banks improve economic efficiency and resource allocations.

The purpose of a bank as monetary intermediary is therefore to improve resource allocations and promote economic efficiency by minimising the risks and costs of transforming risky, privately issued debt into a riskless cash substitute.[61] The centuries long history and continued existence of banks suggests that monetary mechanisms are on aggregate largely effective at adapting to a wide variety of constantly evolving risks

[60] The profit motives of the bank are not included in the model, but the DSGE model of Chapter 4 does incorporate these features.

[61] This description corresponds to the role of banks detailed in early banking literature, such as von Mises (1953 [1912], pp. 263-268), and in contemporary models of banks based on this early bank theory, such as Sissoko (2016).

and costs which must be managed.[62] Nevertheless, failure to manage these risks and costs appropriately will undermine the value of the asset portfolio ($\mathrm{E}(\rho_{K(M)})$) held on a bank's balance sheet relative to its obligations $\left(P'''_{K(M)}\right)$. The actual or even perceived likelihood of a negative net-asset-value—i.e. $P'''_{K(M)} > \mathrm{E}(\rho_{K(M)})$—means that the cash on-demand-and-at-par option will be seen as "in the money", making it rational for deposit holders to exercise the option to convert their deposits into cash, thereby initiating a "run". This guarantee means that banks are, by design, in the business of trust; a breakdown of this trust constitutes the dissolution of this monetary mechanism resulting in a run on the bank. If one bank's assets are in the form of another bank's deposit liabilities, this has the potential to have broader spill over effects and become a broader banking system crisis.[63] In incomplete markets with aggregate uncertainty, monetary intermediation is therefore vulnerable to collapse when trust in underlying asset values is undermined and measures for managing this are inadequate. This vulnerability has policy implications. In terms of the monetary intermediation framework presented, financial stability is enhanced by measures that promote stable, high-quality asset portfolios;[64] measures that increase capital or collateral requirements;[65] measures that enhance risk diversification;[66] and measures that reduce

[62] Merton & Bodie (2005) argue that the "best" financial institutional structure rests on two basic premises. That: (i) financial functions are more stable than financial institutions; and (ii) competition will cause changes in institutional structure to evolve toward greater efficiency in the performance of these financial functions. The Lindy Effect (the theory that the future "life expectancy" of non-perishable things—such as institutions—is proportional to their current age) suggests that bank-like institutions will continue to provide monetary intermediation services for centuries to come (Mandelbrot, 1983; Taleb, 2012).

[63] See the discussions of Wicksell (1936 [1898]), von Mises (1953 [1912], pp. 321-322), and Fisher (1933) for more discussions on the problem of banking crisis. One method of averting such a crisis is to make more cash available in times of such stress—this is the lender of last resort function proposed by Bagehot (1873).

[64] See Dang, Gorton, & Holmström (2015) and Mian & Sufi (2015).

[65] See Holmström & Tirole (1997).

[66] See Sowell (2009) and Brunnermeier & Sannikov (2016).

financial market incompleteness and information imperfections.[67] These policy prescriptions are often given in the context of mitigating moral hazard and other asymmetric information problems. Here, these prescriptions arise without the nature of the friction being explicit, providing a more general motivation that is not restricted to information asymmetries. This discussion is congruent with Gorton & Winton (2003), who conclude that—based on the stylised facts about cross-country banking history—banks do not appear to be *inherently* unstable, and that panics appear instead to be a function of industrial organisation.

The features discussed above describe the essential balance sheet structure of banks. Thus, in moving from a deterministic setting to environments with more uncertainty and market imperfection, the balance sheet structure of a monetary mechanism becomes increasingly recognisably bank-like. The relative proportions of the various balance sheet components necessary to efficiently perform the role of monetary intermediation depend on the broader institutional environment.

DEFINITION 3: A monetary intermediary *is an intermediary whose services provide the functionality of a monetary mechanism.*

PROPOSITION 2: A monetary intermediary will have the characteristics of what in practice constitutes a bank. Banks are monetary mechanisms that transform asset portfolios of debt into liabilities that are a fungible cash substitute. By doing so, banks improve the Pareto efficiency of exchange.

[67] See Mattesini, et al. (2009) and Wallace (2011).

2.12. Further discussion

Continued financial market development, including both widening and deepening of financial markets, along with decreasing transaction costs and improved access to information, has not led to a reduction in bank intermediation as predicted by financial intermediation theory—instead it has increased (Allen & Santomero, 1998; 2001; Scholtens & van Wensveen, 2000; 2003; Coval & Thakor, 2005). Additionally, financial intermediation theory predicts that bank intermediation should play a larger role in developing countries, where information problems are more pronounced, and should be a higher percentage of GDP in emerging economies compared to developed economies. By contrast, financial services are found to be a higher percentage of GDP in more developed economies, and bank intermediation is found to be positively correlated with both GDP per capita and the rate of economic growth (King & Levine, 1993). Why? The discussion in this chapter provides a framework for understanding and explaining these findings. Monetary intermediation theory argues that banks facilitate economic activity by transforming private debt-credit contracts into a substitute for cash. Reductions in financial market imperfections increase the efficacy and efficiency of monetary intermediation, thereby facilitating economic activity. One would therefore expect a positive correlation between monetary intermediation and GDP, and for monetary intermediation services to increase with a reduction in the costs of market imperfections. From this perspective, and contrary to the conventional wisdom of financial intermediation theory, these findings can be interpreted as confirming that monetary intermediation is better suited to environments where financial market frictions, information asymmetries, and other failures are *reduced*.[68]

Monetary intermediation theory also offers alternative answers to other questions raised in the financial intermediation literature. For example, Kashyap, et al. (2002) question why one institution carries out both lending and demand-deposit functions under one roof? Kashyap, et al. argue that the asset and liability sides are synergistic because they can "share" the costly overheads required for providing liquidity. The model presented in this chapter suggests instead that both *bank loans and bank demand-deposits are corresponding and opposing balance sheet entries that account for a single activity*: an exchange between a buyer and a seller. In this case, the dual lending and deposit balance sheet entries are an accounting necessity and reflect the role of banks as monetary intermediaries that transform loans into fungible cash substitute deposits.

---

[68] In response to empirical observations at odds with traditional theoretical priors, Allen & Santomero (1998; 2001) highlight changes to the characteristics of the financial system, suggesting that financial market developments have increased the complexity, and therefore the cost, of *participating* in these markets. In this context, financial intermediaries ameliorate *participation* costs by participating in financial markets on behalf of their stakeholders. A similar case is presented in Scholtens & van Wensveen (2000), who argue that as financial markets change, bank products evolve accordingly. The role of financial intermediation is to provide value creation through risk-sharing and risk-management services that are unavailable through financial markets, but this menu of services continuously changes in response to financial market developments. They find that traditional theory does not account for this dynamic process and is therefore unable to explain why financial intermediation has not decreased in response to financial market developments. These discussions are in accordance with the functional perspective of financial intermediation theory (Merton, 1993; 1995; Merton & Bodie, 1995; 2005). The functional perspective views the interactions of financial markets and financial intermediaries as simultaneously competitive and complementary, and views developments in financial markets as both replacing existing financial intermediation services and at the same time promoting the emergence of new financial intermediation services (Merton, 1993; 1995; Merton & Bodie, 1995; 2005). In essence, both Allen & Santomero and Scholtens & van Wensveen argue that bank product offerings evolve to reduce the costs of a nonstationary set of cost parameters. The discussion presented in this chapter complements these arguments. However, it provides an alternative paradigm for explaining these empirical observations. In terms of this alternative paradigm, bank (monetary) intermediation is more efficient in an institutional environment where markets are more complete and where information asymmetries and other market imperfections are reduced.

Calomiris & Kahn (1991) question why bank liabilities are made payable on-demand-and-at-par? They argue that this feature is designed to overcome issues of moral hazard by imposing bank fragility as an incentive constraint. In contrast to Calomiris & Kahn, monetary intermediation argues that banks are not *designed* to be fragile. Rather, bank liabilities are made payable on-demand-and-at-par in order to function as a cash substitute. As a result, by fixing the price of these liabilities, banks are vulnerable to external shocks from the broader institutional environment, which undermine the value of its assets relative to its liabilities, and make banks susceptible to the large scale exercise of the built-in call option on their deposit liabilities. This can make banks fragile if the measures in place that are designed to buffer these shocks are inadequate. Fragility is therefore a by-product of monetary intermediation, and not a design feature.

2.13. Conclusions

The model presented in this second part of the chapter provides a modelling framework for deriving the characteristics of banks from first principles. Using this framework, I show that bank balance sheet structure and contract characteristics are a direct consequence of both the role of banks as monetary intermediaries, and the broader (inherently uncertain) environment in which banks perform their monetary function of transforming debt into money. In terms of this framework, banks are monetary intermediaries that exist to facilitate the exchange of G&S between buyers and sellers by transforming risky debt-credit contracts (bank assets) into safe money (bank liabilities).

In terms of this framework, banks are not designed to be fragile; rather, bank runs are a failure of banks to adequately account for, and respond to, shocks characterising the broader institutional environment.

Empirical evidence shows that financial market development, reductions in transaction costs, and increased access to information have not coincided with predicted reductions in bank intermediation. The model presented in this chapter reconciles with these findings. I show that more complete and frictionless financial markets improve the efficacy and efficiency of monetary intermediation, and enhance resource allocations and economic efficiency. This provides an explanation for why monetary intermediation has increased with financial market development.

The discussion in this chapter highlights both the role of private debt within the money system, and the role of banks in mitigating the risks and costs associated with using this debt as a medium-of-exchange within this system. The recent financial crisis demonstrated the impact on economic stability when the quality of the private debt underlying the monetary system is in question (Dang, et al., 2015; Mian & Sufi, 2015). The model presented highlights the role of banks as intermediaries in shaping this stability. The crisis has led to a re-evaluation of the policy landscape, and the model presented provides policy guidance in this regard. In terms of the model, bank intermediation is strengthened and financial stability is enhanced by policies that promote high-quality asset portfolios, that increase capital or collateral requirements, that promote asset diversification, that counter pro-cyclical risk build-up, and that reduce financial market incompleteness and information imperfections. These policy prescriptions are often given in the context of limiting moral hazard or asymmetric information problems, but here these prescriptions (while not excluding moral hazard or asymmetric information problems) are far more general and are intended instead to increase the quality of the asset portfolios that underlie the stability of the monetary system.

2.14. Appendices

2.14.1. Appendix 2A: The timing of bank balance sheet entries

In the scenario discussed in Part 1 of the chapter, the buyer, Mr B, is guaranteed a credit facility prior to the sale of the property as a necessary condition for the sale to take place. However, no credit or debit accounting entries occur when the guarantee is granted. The guarantee is simply an endorsement, by the bank, of the buyer's perceived ability to honour his debt obligations. Prior to the transaction taking place, the credit facility will only appear as a note on the bank's balance sheet for prudential purposes.[69] Only on the sale of the property are relevant credit and debit accounting entries made by the bank. These entries record the debt of Mr B as an asset to the bank, and record a corresponding liability deposit entry in the name of Ms S. A record of the increase in the economy's money supply occurs when the bank makes the accounting entry crediting the account of Ms S. The timing of the recording of the increase in money supply has led to confusion concerning the monetary intermediation role of a bank. Two brief examples illustrate this confusion.

[69] This can be readily verified by reviewing bank financial statements compiled in terms of generally accepted accounting principles (GAAP) or international financial reporting standards (IFRS).

First, consider the case where Mr B has access to an overdraft facility. No accounting entries are made at the time the overdraft facility is granted.[70] The overdraft is an endorsement by the bank of Mr B's perceived ability to honour potential future debt obligations. Mr B does not enter into any debt-credit position with the bank when the overdraft facility is approved, and the overdraft is neither recorded as an asset nor as a liability to the bank. Prior to any transactions, the unused overdraft is a conditional commitment and an off-balance-sheet operation. Balance sheet entries occur only after: (i) a cheque has been drawn by Mr B against the overdraft facility; and (ii) the cheque is returned by someone else to a bank to be "deposited". The case of the cheque drawn against an overdraft illustrates that the moment a transaction occurs may not correspond to the moment when the overdraft facility is first approved by the bank, or to the moment when the cheque is returned to be deposited, at which point it is *recorded* on the bank's balance sheets within the banking system. The depositing of the cheque records a transaction that occurred at an *earlier* moment. The cheque itself may circulate widely as a monetary instrument, financing additional transactions before being deposited and recorded by the banking system. This highlights that the *recording* of bank deposits and the *transactions* they facilitate do not necessarily occur at the same moment in time.

Second, consider the case if Mr B is granted a personal loan. If the loan is in the form of a contingent liability, then the description of an overdraft in the paragraph above carries. But if the bank instead chooses to credit a deposit account in the name of Mr B with the value of the loan at the moment the loan is approved, then the money

[70] This can be readily observed in practice, but has also been noted in the literature. For example Cencini (2002, p. 69) notes: "[a]s rightly pointed out by Keynes, unused overdrafts do not 'appear anywhere at all in a bank's statement of its assets and liabilities'".

supply will have seemingly increased without a corresponding transaction. This is an example of a bank deposit being recorded *before* an actual transaction. This describes the setup of Werner's (2014a) study and Jakab & Kumhof's (2015) DSGE model.[71] From Mr B's point of view, this arrangement is only useful if he plans to immediately transfer the bank deposit in order to affect a purchase; because, as noted by Tobin (1963, p. 8): "[b]orrowers do not incur debt in order to hold idle deposits".[72] Accordingly, the credit of Mr B's account would be a reflection of the bank's endorsement of his credit worthiness, to be followed shortly by an actual exchange transaction. If Mr B does not make immediate use of the deposit, then he will be incurring an unnecessary interest expense and would benefit from an overdraft facility instead.

In summary, the mortgage loan discussed involved the *simultaneous* recording of both new bank-money and an underlying transaction. The deposit of a cheque is an example of the recording of new bank-money taking place only *after* the transaction. The example of a loan credited directly to a borrower's deposit account is an example of the recording of new bank-money taking place *before* the transaction. Regardless of the timing, the result is the same: a transaction (involving the exchange of G&S) creates an obligation on the buyer to make future payment. The timing of the recording of a transaction and the broader institutional form of monetary intermediation has historically been driven by a process of regulatory arbitrage, meaning that as specific types of economic activity or groups of economic agents have fallen under specific and

[71] See Chapter 3 for further discussion on these studies.

[72] Additionally, as Graziani (1989, p. 4) writes, "In principle, in a perfectly competitive credit market, no one would borrow money from a bank before a payment comes due. This is the simple consequence of assuming rational behaviour, since there would be no point in borrowing money and paying interest on it while keeping it idle. Money therefore *only comes into existence the moment a payment is made* [original emphasis]. At that moment, in one and the same act, money is created, the borrower becomes a debtor to the bank and the agent receiving a payment becomes the creditor of the same bank."

targeted regulatory regimes, so the techniques of monetary intermediation have adapted to minimise the costs and other impediments imposed by these regulations (Macleod, 1866; Wicksell, 1936 [1898]; de Roover, 1948; Minsky, 2008 [1986]; Greenbaum, et al., 2016; Sissoko, 2016).[73] The rise of the shadow banking phenomenon is a contemporary example of this evolution (Sissoko, 2016). The nature of bank demand-deposits as a social construct that is the product of both the legal and institutional framework, as described by Zelizer (1989) and more recently Cesarano (2014), is therefore evident.

2.14.2. *Appendix 2B: Deriving market equilibrium in the baseline model*

The solution is adapted from Gorton & Pennacchi (1990, p. 54). The market relations are:

$$\text{Demand for } C\text{: } D_C = \psi N P_{K(C)} \qquad \text{(i)}$$

$$\text{Supply of } C\text{: } S_C = \mathrm{S}\big(P_{K(C)}\big).$$

Demand for C is a function of the number of buyers (ψN) and the amount of buyer debt that must be traded per unit of C $\big(P_{K(C)}\big)$. Supply is similarly a function of price. The total quantity C available for supply by sellers (e_{sC}) is:

$$S_C = (1 - \psi) N e_{sC}. \qquad \text{(ii)}$$

At equilibrium, $D_C = S_C$ and *all* K will be exchanged for *all* C,[74] such that if e_{sC} is normalised to unit value, the price per unit of K in terms of C is:

[73] "Regulatory arbitrage refers to innovations designed to circumvent the costs of public regulation" (Greenbaum, et al., 2016, p. 433).

[74] All K will be exchanged for all C because doing so is incentive compatible and maximises agent utility by Equations (v) and (vi).

$$P_{K(C)} = \frac{1-\psi}{\psi} \equiv \overline{P}_{K(C)}. \tag{iii}$$

This intuitive outcome indicates that $P_{K(C)}$ is simply the ratio of buyers to sellers. The price of current goods in units of private debt ($P_{C(K)}$) is the inverse of $P_{K(C)}$ and therefore:

$$P_{C(K)} = \frac{\psi}{1-\psi}. \tag{iv}$$

Without trade, agent utility is simply zero. Incentive compatibility therefore requires that individual agent utility after exchange (U_{jX}) be at least zero. This is proved in Equations (v) and (vi):

$$U_{bX}\big(Q_C\big) = Q_C = \frac{1-\psi}{\psi} e_{bK} > 0 \ \ \forall\, 1 > \psi > 0 \tag{v}$$

$$U_{sX}\big(Q_K\big) = \rho_{K(C)} Q_K = \rho_{K(C)} \tfrac{\psi}{1-\psi} e_{sC} > 0 \quad \forall\, 1 > \psi > 0, \quad \rho_{K(C)} > 0. \tag{vi}$$

Aggregate utility after exchange (TU_X) with e_{sC} = 1 and e_{bK} = 1 (as per the setup assumptions) is:

$$\begin{aligned}
TU_X &= TU_{bX} + TU_{sX} \\
&= (\psi N)\tfrac{1-\psi}{\psi} e_{bK} + \big((1-\psi)N\big)\tfrac{\psi}{1-\psi} \rho_{K(C)} e_{sC} \\
&= (1-\psi)N(1) + \psi N \rho_{K(C)}(1) \\
&= N - \psi N + \psi N \rho_{K(C)} \\
&= N\Big(1 + \psi\big(\rho_{K(C)} - 1\big)\Big) > 0 \quad \forall\, N > 0, \quad 1 > \psi > 0, \quad \rho_{K(C)} > 0.
\end{aligned} \tag{vii}$$

Equation (vii) maximises aggregate buyer utility iff no more C can be acquired by buyers, and maximises aggregate seller utility iff no more K can be acquired by sellers. But at the market clearing equilibrium price, all C and all K are exchanged. Therefore, both buyer and seller utility are maximised and exchange is Pareto efficient.

2.14.3. Appendix 2C: Scarce cash

Pricing in financial markets follows the same derivation as in Appendix 2B. Demand for M is given by $\psi NP_{K(M)}$ and total supply of M is given by $(1-\psi)Ne_{sM}$. With $e_{sM}=\alpha$ and with no intertemporal trade-offs, no information asymmetries, and no trading restrictions, the price of K at which *all* K is exchanged for *all* M is given by:

$$\overline{P}_{K(M)}=\alpha\frac{1-\psi}{\psi}=\alpha\overline{P}_{K(C)}. \tag{viii}$$

From Equation (2.2), Pareto efficient pricing in financial markets is given by $P_{K(M)}=\rho_{K(M)}$. Therefore, consider the following.

(i) If $P_{K(M)}=\rho_{K(M)}>\alpha\frac{1-\psi}{\psi}=\overline{P}_{K(M)}$, then even if the *entire* endowment of M is exchanged for K in financial markets, the quantity of K exchanged will be *less* than the total endowment of buyers' K (because $P_{K(M)}>\overline{P}_{K(M)}$); and

(ii) if $P_{K(C)}=\rho_{K(C)}$ in goods markets, and $P_{K(M)}=\rho_{K(M)}$ in financial markets; then because $P_{K(C)}=P_{K(M)}$ (by construction—a unit of M converts to a unit of C with certainty) there is no arbitrage across goods and financial markets.

(iii) Therefore, if:

$$P_{K(C)}=P_{K(M)}=\rho_{K(M)}>\overline{P}_{K(M)}=\alpha\frac{1-\psi}{\psi}=\alpha\overline{P}_{K(C)} \tag{ix}$$

the quantity of K exchanged for M in financial markets will not exhaust the total amount of buyers' K, *and* the quantity of M held by buyers will not be sufficient to purchase the Pareto optimal quantity of C (given the endowments of K

available).[75] *However*, buyers still have K available to trade for C. In this case, using K as a supplementary medium-of-exchange to M is *Pareto improving*.

2.14.4. *Appendix 2D: Diversification strategies and the timing of diversification*

Pooled diversification could result from six potential strategies. (i) Each seller could achieve a perfectly diversified portfolio of K by exchanging C in exactly equal proportions with each and every buyer.[76] (ii) Sellers could (independent of buyers) coordinate by agreeing prior to transacting (at the beginning of t_0 for example) to pool their K holdings after transacting (at the end of t_0 for example) into a shared mutual portfolio of K. As part of the agreement, each seller would hold a proportional claim against the mutual portfolio. (iii) Sellers could trade portions of their K holdings with other sellers (at the end of t_0 for example) in order to achieve a perfectly diversified portfolio of K.[77] (iv) Buyers could (independent of sellers) reduce ε' by collectively pooling their debt in a mutual-like risk-sharing entity prior to purchasing C (at the beginning of t_0 for example). Each buyer could then trade their individual claims against the mutual entity in exchange for C. (v) Buyers could sell their private debt into a coalition of buyers and thereby become coalition members. Buyers would receive payment from the coalition in the form of a proportional claim against the coalition's

[75] Remember that $\overline{P}_{K(M)}$ is the price at which *all K* is exchanged for *all M* (by definition). But this price is below $P_{K(C)}$ which is the Pareto optimal price of K in goods markets. Therefore if buyers exchanged K at the price $\overline{P}_{K(M)}$ for all M, this would leave buyers with a quantity of M that is less than the quantity required to purchase Pareto optimal C. The total quantity of M available is therefore insufficient to purchase Pareto optimal C.

[76] A perfectly diversified portfolio is derived by holding a portfolio amounting to $\left(N(1-\psi)\right)^{-1}$ units of each and every buyer's endowment of K.

[77] See Tobin (1963), Black (1970), and Fama (1980) as seminal contributions on this strategy.

total portfolio of private debt. Each buyer could then trade their claim against the coalition directly for *C*. (vi) Buyers and sellers could mutually agree prior to transacting (at the beginning of t_0 for example) that buyers will contribute their debt to a portfolio of pooled and diversified private debt, and cede claims against the portfolio to sellers in exchange for *C*. This strategy encapsulates the processes of the preceding five strategies. All six strategies provide sellers with diversification against return variance, increase the price of *K*, and maximise overall social welfare.

Therefore, *sellers* can mitigate risk by constructing diversified and insured portfolios of *K* (Type D1 diversification strategies); alternatively *buyers* have an incentive to increase the price of *K* by diversification through collectively pooling and insuring their debt in a coalition portfolio (Type D2 strategies); or buyers and sellers may coordinate *ex ante* (e.g., at the beginning of t_0; coordination may be implicit due to social convention) to guarantee that both transactions and the mitigation of the risks linked to the transactions occur simultaneously (Type D3 strategies). While these strategies all provide the same risk-sharing outcomes in an Arrow-Debreu (Arrow & Debreu, 1954) setting, they are *functionally* disparate. That is, under D1, buyers purchase G&S in exchange for *risky* financial instruments; whereas under D2 and D3, buyers use *riskless* claims on an *already diversified and insured* pool of debt to purchase G&S. If there is a timing friction such that transaction and risk mitigation are not simultaneous, then:

$$P_{K(M)_{D2/3}} = P_{K(M)_{D1}} + \varepsilon^t \qquad \text{(x)}$$

where $P_{K(M)_{D2/3}}$ and $P_{K(M)_{D1}}$ are the price of buyer capital under corresponding strategy types; and $\varepsilon^t > 0$ is a measure of costs faced due to this timing friction. In this case, D1 (i.e., *ex post* diversification through financial markets) is inefficient compared to D2 and D3. The timing of risk diversification therefore matters in practise and one would not expect to see sellers diversifying *ex post*. In practise and unlike financial markets, bank balance sheets provide a mechanism for *ex ante* coordination.

Chapter 3

Monetary Intermediation, Financial Intermediation, and Money Creation *Ex Nihilo*

"[E]veryone can create money; the problem is to get it *accepted*"

- (Minsky, 2008 [1986], p. 255)

Part 1: Framing the Role of Banks

3.1. Introduction

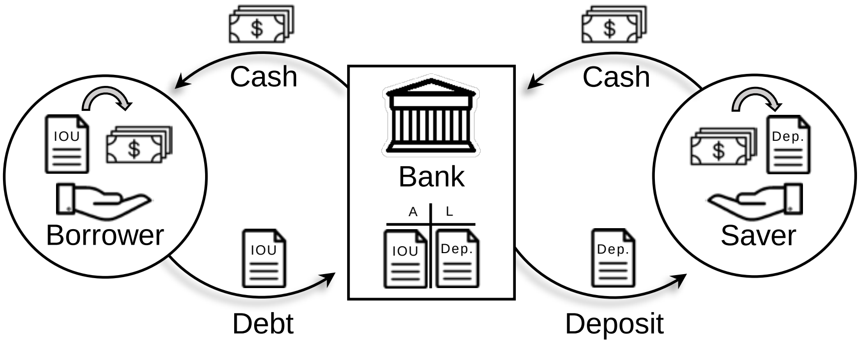
How a bank's role in the monetary system is framed informs how the business of banking is understood and regulated. Financial intermediation theory and theories of money creation *ex nihilo* are two prevalent framings within contemporary bank theory discourse. This first part of the chapter explains how framing the role of banks as monetary intermediaries differs from both the framing of banks as financial intermediaries and the framing of banks as money creators *ex nihilo*.

3.2. Matchmakers, magicians, and mediums

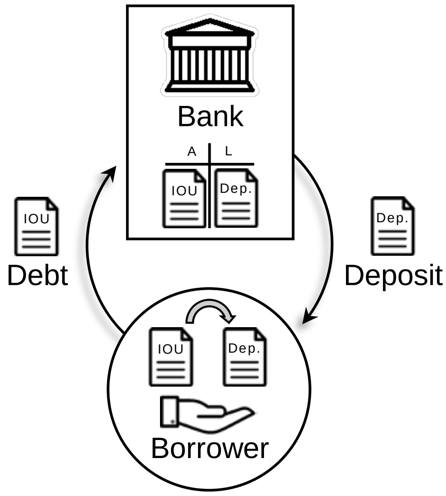
Figure 3-1 illustrates how the framing of a bank's role in the economy differs across financial intermediation theory, the notion of money creation *ex nihilo*, and the theory of monetary intermediation. Each theory places a different relationship at the centre of a bank's economic function. The figure illustrates that the *raison d'être* of banks as monetary intermediaries differs from that described in both financial intermediation theory and endogenous theories of money creation *ex nihilo*.

Figure 3-1: Framing the role of banks in the economy

(i) Financial intermediation theory: Banks are matchmakers that facilitate the transfer of funds from savers to borrowers.



(ii) Money creation *ex nihilo*: Banks are magicians that create money out of thin air for borrowers without any intermediation at all.



(iii) Monetary intermediation theory: Banks are mediums through which buyers and sellers exchange G&S.

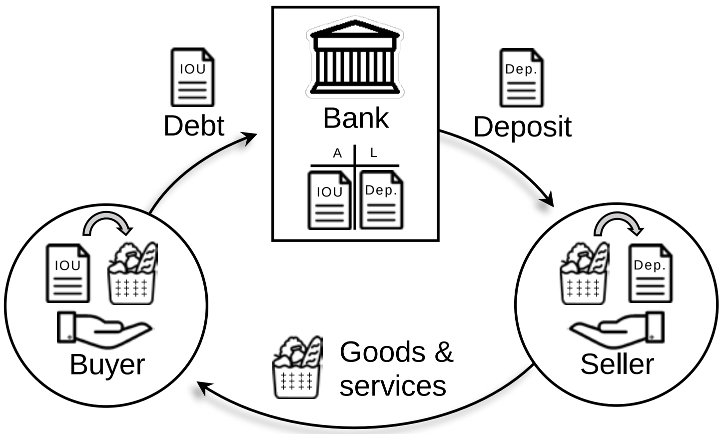


Figure 3-1 illustrates a "matchmaker-magician-medium" paradigm for characterising the different ways in which a bank's role in the economy is framed. In financial intermediation theory, banks are characterised as matchmakers. Banks are middlemen who intermediate the flow of funds between savers (bank liabilities) and borrowers (bank assets). In theories of money creation *ex nihilo*, banks are characterised as magicians. Banks conjure up money out of nothing but "thin air". Banks have no intermediary role at all, as the focus is simply on the relationship between banks and the borrowers who are simultaneously assets and liabilities to the bank. In monetary intermediation theory, banks are mediums; they provide a way for moneyless buyers to transact directly with sellers wanting payment in money. As monetary intermediaries, banks facilitate transactions between buyers (who become the banks' assets) and sellers (who become the banks' liabilities) by transforming the debt of the buyers into money for the sellers.[1]

Financial intermediation theory is largely unconcerned with the role of demand deposits as a medium-of-exchange. Demand deposits are treated first and foremost as a store-of-value asset class held by savers. This is emphasised in the seminal contributions to financial intermediation theory of Tobin (1963) and Fama (1980).[2] Banks intermediate the flow of funds between agents who invest their cash savings in the form of bank liabilities, and agents who borrow these cash savings in the form of bank loans

[1] This framing of a bank's role in the economy therefore corresponds to the role of a bank as understood in the Theory of the Monetary Circuit as described in Graziani (1989, p. 3): "any monetary payment must therefore be a *triangular transaction* [original emphasis], involving at least three agents, the payer, the payee, and the bank."

[2] "The distinction between commercial banks and other financial intermediaries has been too sharply drawn. The differences are of degree, not of kind ... In particular, the differences which do exist have little intrinsically to do with the monetary nature of bank liabilities" (Tobin, 1963, p. 16);
"[Banks] provide transactions services ..., and they provide portfolio management services ...; the concern with banks in macroeconomics centers on their role as portfolio managers" (Fama, 1980, p. 44).

in order to invest in productive economic activity. There is therefore little discussion on the role of bank liabilities as a medium-of-exchange within the monetary system. This view is perhaps surprising for at least two reasons. The first is that bank liabilities constitute the majority share of money supply measures.[3] The second is that the use of bank liabilities as a medium-of-exchange is recognised as a unique and defining characteristic of banks.[4]

According to the theory of money creation *ex nihilo*, the purpose of banks is to transform borrower debt directly into a medium-of-exchange without any role as an intermediary. Banks are said to create money out of nothing when granting a loan to a borrower who becomes simultaneously an asset and a liability to the bank. The role of a bank as *intermediary* between its debtors and its creditors is therefore not a focal issue within the theory of money creation *ex nihilo*. This explanation focuses on a bank's relationship with its debtors (assets), but it says little about the relationship with its creditors (liabilities) once a deposit is transferred. The theory therefore says little about the potential drivers of wider financial crises arising from the liabilities side of a bank's balance sheet (i.e. those factors that determine preferences concerning holding bank deposits).

---

[3] McLeay, et al. (2014a) find that bank deposits constitute 97% of the money supply in circulation in the UK (depending on how on money supply is measured). Similarly, in South Africa, central bank issued notes and coins in circulation constitute only 3% of the South African money supply (M3), or 6% of money supply when excluding all term deposits (M1), as reported in the South African Reserve Bank's Quarterly Bulletin (SARB, 2017, p. S23).

[4] "Demand deposits are securities with special features ... A significant portion of ... the liability side [of a bank's balance sheet] is in the form of demand deposits, securities that have the important property of being a medium of exchange" (Gorton & Winton, 2003, pp. 434-5). "Thus we see that the distinctive function of a bank ... was to issue notes, payable to the bearer on demand, which were to circulate as money" **Invalid source specified.**.

By contrast, monetary intermediation theory emphasises a bank's role as intermediary between its creditors (the sellers of G&S) on the one hand, and its debtors (the buyers of G&S) on the other. As monetary intermediaries the *raison d'être* of banks is to facilitate exchange between these parties by mitigating the risks and costs that would otherwise constrain these transactions. This view puts a bank's role as intermediary at the centre of both the monetary system and at the centre of economic activity more generally. This makes banks special, important, and worthy of specific attention within macroeconomic studies.

3.3. Conclusions

How a bank's role in the economy is framed influences theory developed to explain this role. In financial intermediation theory, banks mitigate the risks and costs of transferring loanable funds between savers and borrowers. According to the theory of money creation *ex nihilo*, banks transform private debt into money without intermediation. In monetary intermediation theory, banks mitigate the risks and costs of buyers and sellers trading G&S on the basis of debt-credit relationships. Regulation and policy are formulated based on the perceived role of banks within the economy. Acknowledging these different framings, testing the validity of these framings, and exploring how the economic implications differ across these framings, is therefore important for shaping an appropriate regulatory environment and for devising effective policy rules for influencing the behaviour of the relevant economic agents.

Part 2: Monetary Intermediation or Money Creation *Ex Nihilo*?

3.4. Introduction

The phrase "banks create money out of nothing" forms part of the popular lexicon. [5] Its popularity has risen in the wake of the 2007/8 financial crisis, fuelling criticism of the banking system by the general public and elected officials alike. But do banks really create money out of nothing? In this second part of this chapter, I argue that the idea of banks creating money out of nothing is a misleading misinterpretation of the role of banks in the economy.

The theory of banks as creators of money *ex nihilo* is often positioned as a revival of earlier bank theory literature. I consider this framing in relation to views expressed in this earlier literature and do not find support for this position. I pay particular focus to the views of Wicksell and Schumpeter as they are often referenced as examples of influential early bank theorists who support this position. I illustrate that both authors viewed banks as monetary intermediaries as opposed to money creators *ex nihilo*.

[5] The phrase 'out of nothing' is also used interchangeably with the phrases 'out of thin air' or '*ex nihilo*' where c*reatio ex nihilo* denotes creation from a state of absolute nothingness, also often used in a theological context in reference to the creation of the universe (Torchia, 1999).

The structure of this second part of the chapter is as follows. Section 3.5 discusses recent examples of the framing of banks as creators of money *ex nihilo*, both in general public discourse and in academic literature. I provide examples of how this framing has influenced public discourse. I highlight shortcomings in this framing and draw attention to recognition of these shortcomings in early bank literature. In Section 3.6, I focus on the views of Wicksell and Schumpeter. This sets up a strong position for understanding the role of banks from the perspective of early banking theory. I highlight how this earlier perspective supports the theory of monetary intermediation as opposed to the notion of money creation out of nothing. This analysis contributes to debate on the role of banks in the economy and the monetary system in particular.

3.5. Creating money *ex nihilo*

3.5.1. Money creation "out of nothing" in public discourse

The notion that banks have the ability to create money out of nothing has raised public ire in the wake of the 2007/8 financial crisis. This unease is understandable as the idea that money can be conjured into existence—by some trick of the banking system, which thereby imposes real obligations on economic agents without any link to real economic activity—is a disconcerting proposition.

The notion that banks create money out of nothing appears in the contemporary public domain, for example in opinion pieces in the Guardian (Williams, 2015); in a speech by Adair Turner, former chairman of the Financial Services Authority (Turner, 2013); and in a report by ratings agency Standard and Poor's (Sheard, 2013). Fuelled by the recent financial crisis, organisations, individuals, and public officials have used public media to call for an end to the current practice of money creation by banks and for an overhaul of the banking system. As examples of the public outcry, Martin Wolf, veteran financial journalist at the *Financial Times*, calls for an end to the practice of money creation by banks (Wolf, 2014). In the US, long serving former senator Ron Paul called for an end to fractional reserve banking and the Federal Reserve System in general (H.R. 1094, 2011). In the UK, the organisation Positive Money continues to campaign against the ability of the current banking system to create money (Positive Money, 2014). In Australia, the online newspaper *The New Daily* writes about the "Orwellian" threat of money creation by banks out of nothing (Stiles, 2017). In South Africa, this view underpinned unsuccessful litigation against the banking system by the organisation The New Economic Rights Alliance (The New Economic Rights Alliance v Absa Bank Limited and others, 2012), who call for an end to money creation by banks. In Switzerland, this view led to a referendum on whether the Swiss banking system should be transformed into a narrow banking system (Dawnay, 2017; Atkins, 2018). This view is therefore not without real world implications. Understanding and clearly communicating the role of banks in the money creation process is therefore important for intelligently informing both public debate and subsequent regulation centred on the role of banks within the financial system.

3.5.2. Money creation "out of nothing" in academic literature

The notion that banks can create money *ex nihilo* is not limited to the public sphere. This idea is found in a broad spectrum of academic literature that influences the public debate. Descriptions of bank-money creation out of nothing appear in relation to Post-Keynesian literature on money (Rochon, 1999a), the Circuitist Theory of money (Rochon, 1999b; Gnos, 2006), works focusing on the theories of Karl Marx (Bellofiore & Realfonzo, 2003; Mouatt, 2011), Modern Money Theory (Wray, 2012), as well as more general literature focused on the theory of banks **Invalid source specified.**. The idea often accompanies discussion on endogenous money creation theories.[6]

McLeay, et al. (2014a; 2014b) published two companion articles in the *Bank of England Quarterly Bulletin* focused on explaining the money creation process and highlighting the fact that this process is not well understood. These articles support the notion that money is created from nothing. McLeay, et al. (2014b, pp. 15-16) write:

> "[V]iewing banks simply as intermediaries ignores the fact that, in reality in the modern economy, commercial banks are the creators of deposit money ... When a bank makes a loan, for example to someone taking out a mortgage to buy a house, ... it credits their bank account with a bank deposit of the size of the mortgage. At that

[6] Endogenous money creation theories form a counterpoint to exogenous money theory, which is influenced by the works of Friedman**Invalid source specified.** and Tobin (1963). The distinction can be traced back to the debate between the currency and the banking schools. According to exogenous theories, money creation centres on the process of banks lending out cash that has been exogenously injected into the banking system. Lending creates a corresponding deposit entry, but this lending is made possible by the pre-existing cash injection (injected for example through a "helicopter drop" of cash). And the process is asymptotically limited by the reserve ratio. This process is précised as "deposits create loans". By contrast, according to endogenous descriptions of money creation *ex nihilo*, money is created when a bank debits a borrowing customer's loan account and simultaneously credits the customer's deposit account with an amount equal to the size of the loan. No prior cash-money savings are required for these accounting entries to take place, and hence there is essentially no (fixed) money multiplier process and no financial intermediation taking place (Moore, 1988; McLeay, et al., 2014b). This process is précised as "loans create deposits".

moment, new money is created. For this reason, some economists have referred to bank deposits as 'fountain pen money', created at the stroke of bankers' pens."

There is a factual inaccuracy in McLeay, et al.'s description as it relates to a mortgage bond,[7] but this error aside, a concern with this explanation (and others like it) is that the act of bank-money creation (i.e., the debiting and crediting of agents' accounts) is disassociated from an underlying exchange of G&S that occurs between economic agents. In the example of a sale of a house subject to a mortgage, as a result of the transaction, the buyer will have a house and the seller will no longer have a house and will instead have money in the form of a bank deposit. The buyer does not receive a house out of nothing, the buyer receives a house in exchange for a debt obligation. The seller does not receive money out of nothing, but in exchange for a house. The intermediary role of the bank transforms the debt of the buyer into money for the seller. In so doing, the bank mitigates various risks and costs borne by these agents in association with this transaction. By separating the granting of credit from the underlying exchange of property, banks play no obvious intermediation role. The descriptions by McLeay, et al. (2014a; 2014b) treat the crediting of a buyer's account and the debiting of a seller's account as two independent activities as opposed to matching accounting entries recording a single activity—a transaction between a buyer

[7] In the case of a mortgage bond, a bank will debit the account of the property buyer (B) and credit the account of the *seller* (S) on conclusion of the sale and transfer. The bank will register a mortgage bond against the house in order for the house to act as collateral for B's debt obligation (the bank's loan asset). For this mortgage bond to be registered, B first has to be the owner of the house, and B will only take ownership of the house once transfer takes place. Until transfer has taken place, the property cannot form collateral against B's loan from the bank, as the property is in fact still owned by S. Hence, in practice, transfer, mortgage bond registration, the debiting of B's account, and the crediting of S's account, all happen simultaneously through the coordination of the parties' various legal agents. Hence, B's account will not be credited with the amount of the loan prior to the transaction, as the bank would be doing so without any mortgage security—and this would (i) not constitute a mortgage bond, and (ii) amount to unsecured lending. Hence the bank never credits B's account in the event of a mortgage loan as stated by McLeay, et al. (2014b).

and a seller. Hence, there is no role for banks as intermediaries within their discussion. Von Mises (1953 [1912], p. 36) pointed out this deficit in the description of how bank credit operates more than a century earlier:

> "Particular attention has been devoted, especially in recent times, to the function of money as a general medium of payment. Indirect exchange divides a single transaction into two separate parts which are connected merely by the ultimate intention of the exchangers to acquire consumption goods. Sale and purchase thus apparently become independent of each other. Furthermore, if the two parties to a sale-and-purchase transaction perform their respective parts of the bargain at different times, that of the seller preceding that of the buyer (purchase on credit), then the settlement of the bargain, or the fulfilment of the seller's part of it (which need not be the same thing), has no obvious connexion with the fulfilment of the buyer's part. The same is true of all other credit transactions, especially of the most important sort of credit transaction – lending. The apparent lack of a connexion between the two parts of the single transaction has been taken as a reason for regarding them as independent proceedings … This is obviously incorrect."

Despite this acknowledgement, the notion that banks create money out of nothing remains pervasive. For example, Jakab & Kumhof (2015, p. 3) describe the process of money creation by banks as follows:

> "[W]henever a bank makes a new loan to a non-bank customer X, it creates a new loan entry in the name of customer X on the asset side of its balance sheet, and it simultaneously creates a new and equal-sized deposit entry, also in the name of customer X, on the liability side of its balance sheet. The bank therefore creates its own funding, deposits, in the act of lending. And because both entries are in the name of customer X, there is no intermediation whatsoever at the moment when a new loan is made."

Jakab & Kumhof (2015) incorporate this concept of money creation *ex nihilo* into a DSGE framework and illustrate that changing the framing of a bank's role in the money creation process has implications for macroeconomic dynamics. They state:

> "Following identical shocks, … [endogenous money creation *ex nihilo*] models predict changes in bank lending that are far larger, happen much faster, and have much larger effects on the real economy than otherwise identical … [benchmark exogenous money] models … Our results suggest that a quantitative investigation of the effects of macroprudential policies using … [money-creation-out-of-nothing] models is likely to yield results that differ significantly from the existing literature." (Jakab & Kumhof, 2015, p. 39)

This statement is pertinent given the increased emphasis on incorporating financial sector mechanics into macroeconomic models following the 2007/8 financial crisis.[8]

Werner (2014a) conducts an empirical field study aimed at demonstrating that banks create money *ex nihilo*. Werner works in conjunction with a small German bank. He is granted a loan by the bank, and the accounting records of the bank's balance sheet are scrutinised during the process. Werner (2014a, p. 16) claims to show that:

> "[I]t can now be said with confidence for the first time – possibly in the 5000 years' history of banking – that it has been empirically demonstrated that … [t]he money supply is created as 'fairy dust' produced by the banks out of thin air … Henceforth, economists need not rely on assertions concerning banks. We now know, based on empirical evidence, why banks are different … it is because they can individually create money out of nothing."

Werner's results are valuable, but his analysis is incomplete because the role of banks as intermediaries is regrettably omitted. This is highlighted by comparing the procedure of

[8] Benes, et al. (2014) and Borio and Disyatat (2011) discuss the theoretical considerations concerning money creation usually omitted from macroeconomic models.

how a loan is granted for the *purposes of the study*, and the bank's *standard procedure for granting a loan* as described in the appendices to Werner's study.[9] In terms of the bank's standard loan procedure, the final step of the process—that is: "Payment of loan at the value date, in exchange for evidence of use of the loan in line with the declared use in the loan application" (p. 18)—does not take place. Werner further states that loans are only ordinarily extended by the bank for an "agreed purpose, as evidence for which a receipt for any purchases undertaken with the loan funds are demanded by the bank" (p. 14). However, Werner's experiment does not involve any intended expenditure; hence, there is no evidence of expenditure, and no receipt for any purchases. The steps usually taken in terms of the bank's standard loan procedure therefore remain incomplete, and there is no connection to an underlying transaction. The role of the bank as intermediary to an exchange transaction is unfortunately omitted from the study. Because Werner will not actually transact on the basis of the loan, steps in the intermediation process which require evaluation and monitoring of Werner's debt obligation are not required. This means the bank's role as risk- and cost-mitigating intermediary is excluded from the setup and results of the study.

Werner's loan does not reflect any credit risk to the bank as the accounting entries are simply reversed at the end of the experiment. There are no costs or risks involved to either Werner or to the bank during this process. At the same time, "the bank waived the actual interest proceeds, in support of the scientific research project" (p. 14). Werner simply becomes equally and simultaneously both creditor and debtor to the bank with a zero net balance sheet position (as shown in the section

[9] The standard loan procedure is described in "Appendix 1 'Sequence of steps for the extension of a loan: Raiffeisenbank Wildenberg e.G.' ".

"Supplementary material 2. The account statement of the borrower") and with zero risk to the bank. Werner's conclusion relating to this process is that banks create money as "fairy dust" and "out of thin air". This interpretation is incomplete because it does not acknowledge the underlying relationship between a bank's debtors and its creditors, nor does it acknowledge the role of the bank as risk- and cost-reducing intermediary. And further, there is a concern with this sort of interpretation that was highlighted by Macleod (1866) more than a century prior.

Macleod's discussion comes in the context of a similar, earlier controversy about whether banks should be permitted to discount bills of *accommodation*, or whether they should be limited to discounting only *real* bills of exchange. In discussing this debate, Macleod describes the discounting of a real bill of exchange,[10] followed by a description of the discounting of an accommodation bill.[11] Macleod highlights the primary distinction between the two as being simply a difference in the timing of the underlying transaction financed by the bill.[12] Macleod points out that the difference in transaction timing had led to a view that accommodation bills were inherently

[10] "In the case … of a bank discounting the bill [of exchange] of the manufacturer A, upon the dealer B, the transaction was already effected upon which it was founded. A had rendered the service to B for which he was to be paid at a future day, before he drew the bill upon him … [T]he bill was given by B, who had got the goods, to A, who had given them, and A had got the money that would be payable to him at the maturity of the bill, advanced to him by the bank [in the form of a bank deposit] on the credit of B's reputation, as well as his own" (Macleod, 1866, pp. 102-103).

[11] "[W]herever an acceptor, drawer, or endorser puts his name upon a bill, and therefore renders himself liable to a holder for value to discharge it, without, as the legal expression is, consideration moving to him, it is an accommodation bill, and the party for whose accommodation it is negotiated is bound by law to provide funds to discharge it at maturity, and also to indemnify the accommodation acceptor, drawer, or indorser [sic], as the case may be, against the consequences of non-payment … [W]hen a bank performs an operation of exactly the same nature, it always receives a high interest rate for the risk it runs, and, when judiciously done, is a very profitable source of income" (Macleod, 1866, pp. 103-104) . The process of a bank endorsing an accommodation bill, would in today's terminology, be the equivalent of a bank debiting a loan account and simultaneously crediting a deposit account both in the name of the *same agent*.

[12] See Chapter 2, Section 2.14.1 for further discussion on differences in timing in relation to the granting of bank credit relative to the underlying transaction in a contemporary context.

problematic, deceptive, and risky *because* of this difference in timing. But Macleod notes that this view gave no consideration to the *quality* of the debt being discounted by the banks' notes. Macleod writes:

> "From the extravagant abuse, however, of such methods of raising capital, accommodation bills have acquired a most discreditable reputation … A great deal has been said and written about the difference between real [bills of exchange] and accommodation bills, and while no terms of admiration are too strong for the first, no terms of vituperation are too severe for the latter … [However, t]he essential distinction between real and accommodation bills is, that one represents a *past* and the other a *future* transaction. But even this is no ground for any preference of one over the other. A transaction that *has been* done may be just as wild, foolish, and absurd as one that *has to be* done … We see, therefore, that the common objections urged against accommodation bills are perfectly futile, and quite wide of mark … The true objection of accommodation paper is of a different nature. When the credit system is carried on duly and properly, and within legal limits, it is the most ingenious method ever devised for promoting commerce, and where it has been cautiously used, has marvellously succeeded in so doing. But it is a very trite remark, that the best things when corrupted become the worst. This is eminently true of paper credit. Universal experience proves that there is nothing so dangerous and pernicious as for individuals to have an undue facility for obtaining credit." (Macleod, 1866, pp. 104-105)

Macleod's point is that, regardless of the type of bill being discounted by a bank, the role of a bank is the same—to mitigate the risks and costs faced by any third party accepting the bank's notes as a means of payment. The view that discounting accommodation bills was inherently different to discounting real bills was a misconception concerning the role of banks as risk- and cost-mitigating intermediaries. Concern was therefore misdirected towards the *timing* of the underlying exchange of

G&S in relation to the issuing of the bill, and *not* to the *quality* of the underlying debt being discounted by the bank.

Thornton (1965 [1802]) highlights this same misunderstanding when discussing the problem of using the term "fictitious" in reference to accommodation bills.[13] Much like how accommodation bills were regarded as *fictitious* in the 18th and 19th centuries (due to a misunderstanding of a bank's intermediary role), so the contemporary framing of banks as creators of money *out of nothing* misinterprets the underlying relationship between a bank's debtors (assets) and its creditors (liabilities). Likewise, the contemporary framing carries an undeserved sense of deception that was also of concern to both Macleod and Thornton. The contemporary sense of deception is highlighted in the following excerpts from de Soto (2009, pp. 198-199):

> [W]ith respect to banks' most important enterprise (accepting deposits while maintaining a fractional reserve), banks actually grant loans from nothing and back them with deposits they also create from nothing. Therefore, rather than credit mediators, they are *ex nihilo* creators of credit … the chief source of their profit is the *ex nihilo* creation of money, which provides banks with financing indefinitely. Banks appropriate these funds for their own benefit and charge interest on them to boot. In short, bankers create money from nothing, loan it and require that it be returned with interest."

[13] "The supposition that real bills represent property, and that fictitious [accommodation] bills do not, seems, therefore, to be one by which more than justice is done to one of these species of bills, and something less than justice to the other … A fictitious bill, or bill of accommodation, is evidently, in substance, the same as any common promissory note … the epithet fictitious has been cast upon it; an epithet which has seemed to countenance the confused and mistaken notion, that there is something altogether false and delusive in the nature of a certain part … of the paper" (Thornton, 1965 [1802], p. 87).

The sense of injustice in de Soto's passage resonates with the usury laws of medieval Europe;[14] and as Ferguson (2009, p. 2) observes:

> "Throughout the history of Western civilization, there has been a recurrent hostility to finance and financiers, rooted in the idea that those who make their living from lending money are somehow parasitical on the 'real' economic activities of agriculture and manufacturing."

The framing of banks as creators of money *out of nothing* has an implicit undertone that banks *do nothing* to add value in this process. To the contrary, banks play a crucial role in this process as intermediaries that mitigate the risks and costs of transacting based on debt-credit agreements. Failure to appreciate this point has fuelled indignation and hostility in contemporary public discourse.

3.6. The views of Wicksell and Schumpeter

I have argued and presented evidence that the view of banks as creators of money out of nothing is widely held. Academic literature often references the works of Wicksell and Schumpeter in support of this view. In certain instances, the support of these influential authors is implied by the context of the discussion—see for example Arena & Festré (1996), Werner (2014a; 2014b), Daastøl (2015), Jakab & Kumhof (2015), and Turner (2015). In other instances, this interpretation is explicitly stated—see for example Kregel (1986), Rochon (1999a), Fontana (2007), and Bellofiore (2009). Both Wicksell and Schumpeter are pioneers and important contributors to the discussion on the role of bank credit in the real economy, and this gives their ideas considerable weight.

[14] See for example the discussion in Ekelund, Jr., et al. (1989).

However, in this section I argue that the idea that banks can create money *ex nihilo* is an inaccurate reflection of the views expressed in the (translated) works of both Wicksell and Schumpeter. I argue that neither Wicksell nor Schumpeter supported the view that banks create money out of "thin air". Rather, both explain that the guarantees provided by the banking system enable claims (bank demand-deposit liabilities) against private debt-credit contracts (bank loan assets) to function as a generally accepted medium-of-exchange (money).

In this section, I illustrate that both authors shared a common understanding of a bank's monetary intermediation role. The views of Wicksell and Schumpeter are also consistent with other early bank theory contributors such as Thornton (1965 [1802]) and Macleod (1866; 1889; 1906).[15] Wicksell's focus is on explaining monetary phenomena, whereas Schumpeter's focus is on explaining economic growth within the context of debate taking place at the time. Nevertheless, both authors acknowledge the importance of the guarantees provided by the banking system in enabling claims against privately issued debt based financial instruments to function as money. Wicksell provides a detailed account of this role in the context of explaining the monetary phenomena of interest and prices, whereas Schumpeter provides extensive discussions on the role of bank credit as a determinant of economic growth. Festré (2002) traces the likely influence of Wicksell on Schumpeter's views on banking; Arena & Festré (2005) argue that Schumpeter's discussions on money are an extension of Wicksell's work; and Festré & Nasica (2009) analyse Schumpeter's understanding of money from an institutional perspective. Adding to this, I show that Wicksell's explanation of the

[15] See Chapter 1, Section 1.4 for a brief review of this earlier literature.

banking system's role in converting private debt-credit contracts into a cash substitute provides insights into Schumpeter's discussion on the role of credit and banking in financing exchange through money creation. It is therefore useful to jointly explore the views of both authors, in particular for understanding Schumpeter's views on money creation. The analysis in this section suggests that the notion of banks as creators of money out of nothing is misleading shorthand for a more complex money creating process as understood by both Wicksell and Schumpeter.

The discussion in this section also contributes to the practice of building ties between the Swedish and Austrian traditions.[16] This practice examines the influences among Wicksell, Schumpeter, von Mises, Hayek, and Keynes—and by extension the schools of thought influenced by the so-called Wicksell connection (Leijonhufvud, 1979; Fontana, 2006; Goodspeed, 2012). The discussion in this part of the chapter is therefore pertinent in light of concerns about the role of banks in undermining broader financial stability, which, while central to these schools of the Wicksellian tradition, has only recently re-emerged as central to contemporary policy and regulatory debate in the wake of the recent financial crisis (Laidler, 2009).

3.6.1. Wicksell

Knut Wicksell's *Interest and Prices* (1936 [1898]) contains an extensive discussion on the velocity of circulation of money. Wicksell recognises that a system of organised credit is provided via the banking system and gives considered attention to the role of banks

[16] See de Boyer (1985), Holtfrerich (1989), Flanders (1990), Trautwein (1996), Graziani (2003), and Hagemann (2015) as examples of this practice.

within the monetary system. Wicksell distinguishes between money issued by the monetary authorities (occasionally referred to as cash), and a system of money *substitutes*. These money substitutes arise from the discounting of bills of exchange, and are in the form of bankers' notes, cheques, and other financial instruments that are issued by institutions other than the government's central monetary authority. These money substitutes function as money through the centralisation of lending activities provided by the banking system, and their use serves to increase the velocity of circulation of cash. As stated by Wicksell:

> "In a developed credit economy ... the velocity of circulation is capable of being increased more or less at will. The available means can be placed in two main categories: the transfer of claims (the use of bills of exchange) and the centralisation of lending in monetary institutions; and in addition there is the combination of the two classes which characterises the modern system of banking and bourses (discounting, the use of cheques, clearings, notes, etc.)." (Wicksell, 1936 [1898], pp. 62-63)

Wicksell's observation that the circulation can be increased "more or less at will" is a point to which I will later return.

Wicksell (1936 [1898]) describes the simple bill of exchange as an inadequate substitute for cash-money—it lacks general acceptability as it relies on the quality of the relationship between the transacting parties and suffers from the dependability of the initial issuer. But an informal system of concurrent endorsement enabled these promissory notes to function as practical money substitutes:

> "We have seen how simple merchandise credit and the lending of money from one individual to another, while not capable of providing a definite substitute for money, do give rise to an increase in the velocity of circulation … [However, it] is of course usual for the acceptability of such a promissory note to be further strengthened through each endorser undertaking by means of his endorsement a subsidiary liability. A payment made by means of a bill of exchange is therefore not final, for if the bill fails to be met the money can be demanded from the endorser. It is precisely for this reason that bills provide a great source of strength to the credit system (this was particularly the case in earlier times) … [T]he subsidiary liabilities which are incurred as a result of endorsement are completely discharged only when the bill is finally paid. It might thus be imagined that the sum of money in question had actually been passed along the whole line of intermediaries … It turns out therefore that bills merely provide once again a method of causing a virtual acceleration in the circulation of money. Their superiority over simple merchandise credit here lies in the greater security and negotiability of the promissory documents, which can be included in cash holdings and in reserves almost in the same way as [cash] money … and they can thus actually dispense with [cash] money." (Wicksell, 1936 [1898], pp. 62-64)

This process of circulating endorsement bills of exchange as described by Wicksell was, however, supplanted by the centralised arrangement offered by the modern banking system. And so Wicksell writes:

> "It is only on account of their relative inconvenience, and for similar reasons … that this *original* use of bills has gradually become less prevalent since banking was perfected … "[T]he use of bills might, under favourable conditions, have developed so far as to have dispensed almost completely with the use of [cash] money … [but a] more powerful influence in this direction has been exerted … [by] the development of the banking system … There is no real need for any [cash] money at all if a payment between two customers can be accomplished by simply transferring the appropriate sum of [cash] money *in the books of the bank*. It can be written off the account of [or debited to] the debtor (the buyer) and credited to the account of the creditor (the seller)." (Wicksell, 1936 [1898], pp. 65-68)

The role of a bank in converting private debt-credit contracts between buyers and sellers into a cash substitute is explicit. The banking system is further described as serving an intermediary role between the transacting parties. This intermediary role facilitates exchange through the double entry bookkeeping of the banks, and as a result of the promissory arrangements provided through the banking system, provides a practical money substitute. Wicksell's understanding of the role of a bank is thus described as being an intermediary of underlying transactions between buyers and sellers transacting in terms of debtor-creditor relationships. This understanding is again conveyed where Wicksell writes:

> "The actual exchange of commodities proceeds very simply. The buyer draws a cheque ... for the appropriate sum, and the seller cashes the cheque, the sum being thus credited to him by the Bank. But within a short space of time goods must be paid for by goods ... [and a] certain interval will ... elapse between the sale of one lot of goods and the purchase of another equivalent lot. During this time, the sellers are in reality extending credit to the buyers to the amount of the sum in question... although on the surface the payment has the appearance of being immediate. This is brought about as a result of the facilities and the guarantees provided by the Bank." (Wicksell, 1936 [1898], pp. 71-80)

Wicksell is therefore unambiguous in acknowledging that underpinning the debt-credit based system of cash-money substitutes provided by banks, is the exchange of G&S between buyers and sellers entering into debtor-creditor relationships—the process of which is facilitated by a bank. The value of the G&S exchanged determines the value of cash substitute created, and the debt of the buyer underwrites the value of this cash substitute. The banking system is described as largely analogous to credit systems that exist even without the facilitation of a formal banking system. To the extent that money is said to be created, it is not cash-money as issued by the monetary authorities, but something which functions as a cash substitute and that is recorded within the banking system as a credit (or debit) in favour of the seller (or buyer). It is the bank's liability which functions as a cash substitute (also commonly called bank-money) and which is being referred to when it is said that banks create money. It is thus clear that Wicksell did not hold that banks create bank-money out of thin air; instead the creation of bank-money is linked to an underlying exchange transaction, and the value of the money created is determined by the value of the goods being exchanged.

Two principles governing this money creation process are evident: the first is that the exchange process ultimately results in a financial instrument that is held in the possession of the seller in lieu of real G&S; the second is that for this financial instrument to function effectively as a medium-of-exchange, it must be generally accepted in any further transactions. Wicksell notes that a promissory note issued by a buyer does not readily achieve general acceptability:

> "It is clear that simple credit, as between individuals, has only a very limited influence as a substitute for money tending to accelerate the velocity of circulation. Here the discharge of claims or debts as well as the exchange or transfer of claims is only the exception. Credit for goods is certainly very common between individuals, but it is combined, especially over longer periods, with difficulties and risks ... Organized credit tends to reduce risks by spreading them over a wider area; the subjective element of risk disappears in proportion as the wealth which affords the guarantee is great in relation to the amount at stake, so that only the mathematical risk remains. In this way, and also through the centralization of credit facilities, it helps to make loan transactions safe and convenient. The very documentation of credit transactions in the form of credit instruments, their transfer to others and eventual conversion into claims for payment, valid in the hands of each holder, creates a powerful organization which it has required thousands of years to develop." (Wicksell, 1978 [1906], pp. 70-72)

It is the banking system's guarantee to take over the default risk of the buyers and provide the sellers with tranches of cash *on demand* that makes these cash substitutes (accounting entries recorded in the books of the banks) generally acceptable. The bank's guarantee bypasses the need for successive endorsement by offering an overarching guarantee of convertibility which encourages acceptability. The guarantee offered by the banking system means that these bank liabilities are accepted as payment in subsequent exchanges.

To the extent that money can be said to be created by the bank (understood to be a cash substitute), it is clear that the role of the bank in this process is to reduce the risks and costs associated with the underlying transaction. This description of the role of banks in mitigating risk and reducing transaction costs is consistent with financial intermediation theory (Greenbaum, et al., 2016). However, Wicksell's description focuses on a buyer-seller relationship as opposed to a borrower-saver relationship. The value of the money reflected in the books of the bank is determined by the value of the transaction agreed upon between the buyer and the seller conducting the underlying exchange. This value is determined in the same manner that a bill of exchange is created and obtains its value from the value of an underlying transaction. Without underlying transactions, there would be no need for bank-money. The bank as intermediator is one of three parties involved in the money creation process, although *prima facie*, the involvement of the buyer and seller may not be obvious. Wicksell's description illustrates that the primary role of the banking system is to facilitate exchange by providing financial instruments that are generally accepted as a medium-of-exchange, and that function as cash substitutes. In the context of Wicksell's writings, the process of money creation by banks should be properly understood to mean the process by which: (i) banks issue claims (demand deposits) against a portfolio of financial instruments (private debt-credit contracts); (ii) the value of bank deposits is underwritten by the value of this portfolio; and (iii) banks provide a guarantee of convertibility of deposits into cash-on-demand-and-at-par. The notion that this money is created *ex nihilo* is a misrepresentation of Wicksell's view: bank deposits (bank-money) are not created out of nothing, they are created as a result of an underlying value-for-value exchange transaction between buyers and sellers, and they derive their value from the value of the G&S exchanged.

Wicksell describes cash substitutes as being created at will, which, in the context of the discussion, suggests they are created at the will and discretion of the transacting parties. I can find no evidence to conjecture that this should be interpreted as meaning creation out of nothing. Nevertheless Wicksell does at times appear to offer support to the notion that banks create money out of nothing. Wicksell writes:

> "[M]odern forms of credit … almost always imply the mediation of some bank or professional money-lender. The banks in their lending business are not only not limited by their own capital … they themselves create the money required, or, what is the same thing, they accelerate *ad libitum* the rapidity of the circulation of money. The sum borrowed today in order to buy commodities is placed by the seller of the goods on his account at the same bank or some other bank." (Wicksell, 1907, p. 214)

In this passage, banks are once again acknowledged to be facilitators of transactions between buyers and sellers transacting on the basis of debt-credit contracts. But Wicksell states that banks are able to create money, and then contradictorily states that this act is synonymous with increasing the velocity of money. Taken out of the context of Wicksell's previous work, the statement is contradictory as *money* and *velocity* are distinct and separable concepts (as understood by Wicksell himself). However, in the context of Wicksell's earlier work, the meaning is clear: banks do not create cash—the stuff of monetary authorities—because they deal in cash substitutes. And as discussed, Wicksell earlier acknowledged that banks play an integral role in enabling cash substitutes to operate as such. By Wicksell's earlier explanation, cash substitutes serve as if they are cash and in so doing, have the same effect as increasing the circulation of the issued cash-money. The apparent contradiction in Wicksell's later statement is therefore clarified when viewed in the context of his previous work. Moreover, the banks are said to act *ad libitum*, which is in accordance with Wicksell's previous statement that money substitutes are created "at will". Again, I take this to mean that banks, like any economic agents entering into a transaction involving a debt-credit arrangement, are free to engage at their own will. There is no indication that this should be interpreted as meaning that banks create money *ex nihilo*.

Wicksell describes a credit system that is driven by an underlying exchange of G&S with the accompanying accounting entries and bank guarantees functioning in lieu of cash. Interpretations suggesting Wicksell promoted the contemporary notion that banks create money *ex nihilo* is a misreading of Wicksell. I can find no other passages that would contradict this statement, nor any more relevant passages.

3.6.2. Schumpeter

The role of a bank in the money creation process—as intermediator between parties conducting intertemporal exchange transactions—is common to both the works of both Schumpeter and Wicksell. Additionally, the analysis of Wicksell's explanation in the previous section provides useful insight into Schumpeter's explanation of the money creation process. Schumpeter's explanation distinguishes between the role of debt-credit based bank-money creation as the definitive tool for entrepreneurial growth, as well as a mechanism for speculatively driven inflation (Schumpeter, 1956). I argue that while Schumpeter describes the money creation process as driven by an underlying exchange transaction (as does Wicksell), excerpts from Schumpeter's writings when taken out of context and at face value are potentially misleading. For example, Schumpeter states:

> "[From] whence come the sums needed to purchase the means of production necessary for the new combinations if the individual concerned does not happen to have them? … This method of obtaining money is the creation of purchasing power by banks … It is always a question, not of transforming purchasing power which already exists in someone's possession, but of the creation of new purchasing power out of nothing." (Schumpeter, 1983 [1934], pp. 72-73)

By exploring the context of this statement, I show that to interpret it in the contemporary sense that "banks create money *ex nihilo*" is unfounded. Rather, Schumpeter uses the terminology "out of nothing" to mean a lack of real resources *currently* available, but which are *expected* to be available in the future, and which are therefore contracted against in the present, thus constituting an intertemporal exchange of real G&S. This understanding of the origin of purchasing power created through bank intermediation is consistent with Wicksell's explanation, as well as other early writers such as Macleod (1866; 1889; 1906).

It is useful to acknowledge that Schumpeter frames his discussion on the role of banks in the money creation process within the broader context of economic growth. In Schumpeter's story, economic growth is driven by entrepreneurs who wish to redirect real resources towards their own productive aims:

> "The fundamental notion that the essence of economic development consists in a *different* [original emphasis] employment of *existing* [original emphasis] services of labor and land leads us to the statement that the carrying out of new combinations takes place through the withdrawal of services of labor and land from their previous employment … As a rule the new combinations must draw the necessary means of production from some old combinations … The carrying out of new combinations means, therefore, simply the different employment of the economic system's existing supplies of productive means … Procuring the means of production … [is not a problem for entrepreneurs if they] … have the necessary productive means or can get them in exchange for others which they have … In the contrary … [the entrepreneur] must resort to credit if he wishes to carry out a new combination, which cannot … be financed by returns from previous production … [C]redit is primarily necessary to new combinations … [I]n carrying out new combinations, "financing" as a special act is fundamentally necessary … [D]etaching productive means from the circular flow and allotting them to new combinations … is done by credit, by means of which one who wishes to carry out new combinations outbids the producers in the circular flow in the market for required means of production." (Schumpeter, 1983 [1934], pp. 65-71)

The need for credit as a constraint faced by Schumpeter's entrepreneur is central to Schumpeter's discussion. Schumpeter identifies both "current credit within the circular flow" (which is essentially "certificates for existing goods and past services") and "the credit phenomenon" (which involves the creation of purchasing power that is not backed by *existing* goods or *past* services) as means with which to withdraw resources from their current employment. It is Schumpeter's "credit phenomenon" that is the focus of this discussion. Schumpeter's discussion centres on entrepreneurs being able to secure the use of real resources through credit facilities without having any *currently existing* real resources to offer in exchange. The ability to rely on credit in order to secure access to the stock of currently existing real resources is necessary for the entrepreneur who is only able to offer *future* real resources in exchange for already existing current real resources.

Schumpeter describes the entire monetary system as a system of debts and credits which are ultimately settled by the exchange of real goods against other real goods:

> "The monetary circulation is, in its nature and main function … nothing but a clearing system … Money is a 'claim ticket and receipt voucher' … only commodities are exchanged against commodities in the economy; money is always only an intermediate link without independent significance." (Schumpeter, 1956, pp. 155-171)

Access to credit facilities provides the entrepreneur (a buyer) with current purchasing power, which is to say, the ability to secure the use of currently existing real resources. This credit purchasing power serves as money, and the entrepreneur relies on the facilities of the banking system in order to achieve this purchasing power:

> "These credit means of payment, that is means of payment which are created for the purpose and by the act of giving credit, serve as ready money in trade, partly directly, partly because they can be converted immediately into ready [cash] money … The banker, therefore, is not so much primarily a middleman in the commodity "purchasing power" as a *producer* [original emphasis] of this commodity." (Schumpeter, 1983 [1934], pp. 73-74)

> “That this creation of means of payment centres in the banks and constitutes their fundamental function, we find agrees with the prevailing conception. The creation of money by the banks establishing claims against themselves, which is described by Adam Smith and indeed by still earlier authors … has become a common place to-day; whereupon I hasten to add that for our purposes it is all the same whether or not one regards the expression ‘creation of money’ as theoretically correct.” (Schumpeter, 1983 [1934], p. 98)

Schumpeter describes banks as the producers of money, but additionally acknowledges that the terminology of money creation is in need of clarity, without attempting to provide such clarity. To consider the bank as producer of money was sufficient for Schumpeter’s purposes, while acknowledging that the details of the relationship may not have been accurately represented by the statement “creation of money”. Within this context, Schumpeter feels it was justified to consider the means of payment enabled through the credit system as “money”, as he states:

> “It is well known too that the old question whether … bank credits are money has been answered affirmatively by the best writers.” (Schumpeter, 1983 [1934], p. 97)

And later Schumpeter writes:

> “There surely must be other ways of expressing these practices than by calling these bank notes embodiments of velocity of circulation – a velocity so great that it enables a thing to be in different places at the same time … Therefore it may, perhaps, seem more natural to say that bankers increase not the velocity but the quantity of money – or of those means of payment that, within limits, serve as well as money if one wishes to reserve this term for coin or coin and government paper … Banks are no longer said to ‘lend their deposits’ or ‘other people’s money,’ but to ‘create’ deposits or bank notes: they appear to manufacture money rather than to increase its velocity or to act – which is a completely unrealistic idea – on behalf of their depositors.” (Schumpeter, 2006 [1954], p. 304)

There are two aspects to the notion that banks create money out of nothing which Schumpeter addresses directly. The first aspect is that it is the banks themselves which create money, and the second is that they do so out of nothing. In addressing the first aspect, Schumpeter (consistent with Wicksell) makes clear that there is a connection between the granting of credit and an underlying exchange transaction:

> "[T]he entrepreneur [cannot] himself borrow produced means of production … he can buy them; for this, however, he again needs purchasing power. But he cannot simply borrow them, for they are needed for the purposes for which they were produced, and the possessor cannot and will not wait for his return – which the entrepreneur could indeed reimburse him for, but only later – also can and will bear no risk. If someone nevertheless does this, then two transactions occur, a purchase and an extension of credit. Both are … two legally distinct parts of one and the same process." (Schumpeter, 1983 [1934], p. 96)

Further, Schumpeter acknowledges that all manner of documents acknowledging credit-debt arrangements may serve as means of payment, and that as long as they broadly circulate—thereby providing purchasing power—they are money, regardless of whether they originate from within the banking system or not:

> "That all forms of credit, from the bank-note to book-credits, are essentially the same thing, and that in all these forms credit increases the means of payment, is also held to be uncontroversial." (Schumpeter, 1983 [1934], p. 99)

> "This explains the manner in which the creation of purchasing power works … The external form of the credit instrument is quite irrelevant … a bill, which does not replace existing money, and which is not based upon goods already produced, is of the same character, if it actually circulates … And all other forms of credit instruments, even simple credits in the book of a bank may be considered from the same point of view." (Schumpeter, 1983 [1934], p. 109)

> "[W]e have to include in the concept of money credit instruments and claim titles of all kinds, to the extent that they perform the role of money ... there is no alternative to including all credit instruments – claim titles of all kinds – to the extent that they are in actual, though not necessarily general circulation. Such instruments differ only slightly from bank notes and current accounts, particularly in the essentially irrelevant respect that they are usually not issued by the regular originators of most money creation, the banking system." (Schumpeter, 1956, p. 171)

In light of the discussion pertaining to Wicksell, these statements clarify that in the process of granting purchasing power by means of credit arrangements, the underlying nature of the relationship between the bank and the transacting parties described by Schumpeter corresponds to that described by Wicksell. That Schumpeter's purchasing power comes about due to the facilitation of the bank, speaks to the importance of the bank's intermediation function in enabling general acceptability, as explained in detail by Wicksell. And while the bank may not be necessary for this general acceptability (as alternative methods of risk reduction may also yield acceptability), it is clear that, to the extent that the bank is involved in the process, the bank is more than a mere middleman—a view that is elaborated on by Wicksell. The role of the bank in the money creation process can therefore be seen as one whereby the bank enables claims (bank deposits) against private debt-credit contracts (bank assets) to function as a *generally accepted* medium-of-exchange cash substitute. The bank does this by guaranteeing convertibility of these claims into cash-on-demand-and-at-par.

With regards to the second aspect of the notion that banks create money out of nothing (being that money is created *out of nothing*), Schumpeter clarifies his position. Schumpeter explains that purchasing power results from the value of the future goods that are pledged in exchange for current goods. Schumpeter clarifies that the term "nothing" refers to a lack of G&S *currently in existence*. I find that Schumpeter's money, like Wicksell's, is the result of an *intertemporal* exchange of property:

> "It is perfectly clear that purchasing power is created to which … no new goods correspond … [and] in the absence of past productive services could only be described as certificates of future services or of goods yet to be produced." (Schumpeter, 1983 [1934], p. 101)

> "The entrepreneur must not only legally repay money to his banker, but he must also economically repay commodities to the reservoir of goods – the equivalent of borrowed productive means; or, as we have expressed it, he must ultimately fulfil the condition upon which goods may normally be taken out of the social stream." (Schumpeter, 1983 [1934], p. 110)

> "[I]t appears to me correct to say: although not by existing goods, the quantity of new purchasing power that it is possible to create is supported by future goods, and, to repeat, by *future* goods at *present* prices." (Schumpeter, 1983 [1934], p. 115)

To understand Schumpeter's "out of nothing" as *ex nihilo* is to confuse the notion of *possession* with that of *property*,[17] as the money created is the record of a transfer of the property of the future goods exchanged for that of the current goods (Macleod, 1866; 1889; Heinsohn & Steiger, 2000).[18,19] Debt-credit financial contracts are created in the

[17] "As a title secured by property, and with all due right without intrinsical [sic] value, money can appear to be created *ex nihilo* only from the viewpoint of possession … However, money is never created *ex nihilo* from the viewpoint of property, which must always exist before money can come into existence. Starting with property, one can eventually complete Joseph Schumpeter's unfinished critique of the real bills doctrine … The observation that the creation of money is not limited by the value of already-produced goods is correct." (Heinsohn & Steiger, 2000, p. 79)

[18] "[T]here is also a second species of Property, which is equally the subject of exchange … we may have a Property, or Right, to a thing, which is not even yet in existence. And this species of Property is termed in English law, Incorporeal Property, because it is Property separated from any specific corpus." (Macleod, 1866, p. 12)

[19] "There is an essential distinction between the Right of Property and the mere Right of Possession … Some of the most subtle and important doctrines in Economics are based entirely on the distinction between Right of Property and Right of Possession." (Macleod, 1889, p. 35)

act of intertemporal exchange as the medium through which property is exchanged.[20] The guarantees provided through the banking system enable claims against these contracts to function as cash substitutes.[21] Schumpeter points out that this money creation process need not be the special reserve of the banking system alone, and that the money created is not created out of nothing, but derives value from the value of the property being exchanged:

> "As a rule, [if] an ordinary bill of exchange does not so serve [as a generally accepted means of payment]; then it is not money, and belongs to the demand side of the money market. Sometimes, however, certain classes of them do; then, according to this view, they are money and form part of the supply on the money market. Bank notes and checking deposits eminently do what money does; hence they are money. Thus credit instruments, or some of them, intrude into the monetary system; and, by the same token, money in turn is but a credit instrument, a claim to the only final means of payment, the consumers' good. By now this theory – which of course is capable of taking many forms and stands in need of many elaborations – may be said to prevail … Credit as a creator of money!" (Schumpeter, 2006 [1954], p. 305)

In this final extract, Schumpeter removes reference to banks from the money creation process entirely, replacing banks with the concept of credit. And so it is apparent that Schumpeter's statement of money creation by banks "out of nothing" is shorthand for a more general description of the role played by banks in the transformation of credit into a cash substitute in order to facilitate the exchange of real G&S.

[20] "The true function of Credit is now apparent. It is a very common idea that Credit is the goods which are 'lent': or the transfer of them. Such ideas are utterly erroneous. We have shown that a Credit is the Present Right to a Future Payment. *And the true function of Credit is to bring into commerce the Present Values of Future Profits* [original emphasis]." (Macleod, 1889, p. 80)

[21] "Demand deposits are securities with special features. They can be denominated in any amount; they can be put to the bank at par (i.e., redeemed at face value) in exchange for currency. These features allow demand deposits to act as a medium of exchange." (Gorton & Winton, 2003, p. 1)

3.7. Conclusions

Detailed expositions of a bank's role in the money creation process, and the fundamental link between this process and underlying real economic transactions, are common to early banking theory. This early theory accompanied the birth of the modern banking system in the 19th and early 20th centuries when bank demand-deposits were emerging as a common and generally accepted means of payment. Yet the connection between bank demand-deposits and underlying exchange transactions is absent from contemporary explanations focused on banks as money creators *ex nihilo*.

Emphasising the connection between an underlying exchange transaction and the process of bank-money creation highlights a bank's role in the money creation process as a cost- and risk-mitigating intermediary. In terms of this process, money is not created *ex nihilo,* but as a result of an underlying value-for-value exchange transaction within a broader legal and institutional framework. As described by both Wicksell and Schumpeter, the role of the banking system is to act as intermediary to transactions between buyers and sellers of real G&S: banks enable derivative claims (bank demand-deposit liabilities) against privately issued debt-credit contracts (the bank's loan asset portfolio) to function as a generally accepted medium-of-exchange. Literature framed by the notion that bank-money is created *ex nihilo* omits this intermediary role. This detracts from the quality of debate aimed at regulating the role of banks in the monetary system today.

The notion that banks create money *ex nihilo* invokes a misleading caricature of a bank's role in the process of money creation. The framing of a bank's role in the money creation process is important because it influences public opinion and informs debate, which ultimately frames and shapes regulation. As reflected in works of early bank theorists, bank-money is at its roots a record of exchange based on a debt-credit relationship between transacting parties. A statement of a bank's role in the money creation process as reflected in this early bank literature could therefore read as follows: *banks intermediate transactions based on debt-credit relationships between buyers and sellers; on behalf of these agents, banks manage the idiosyncratic risks and costs associated with the debts of the buyers, and provide sellers with guarantees of cash-on-demand-and-at-par. As a result of these services, the liabilities of banks function directly as a medium-of-exchange and a substitute for cash*. This statement describes the role of banks as monetary intermediaries. While this statement reflects a view in early bank theory, this statement is still accurate and relevant today; and it highlights the relevance of this literature to understanding today's banking environment, despite the regulatory and institutional changes that have taken place since.

The function of banks as monetary intermediaries is to transform risky private debt-credit contracts into a riskless cash substitute. This function is not readily reflected in the phrase "banks create money out of nothing". Regulating the banking system's ability to *create money from nothing* is a different policy framing to that of regulating institutions that *intermediate transactions* between buyers and sellers.

Part 3: Monetary and Financial Intermediation

3.8. Introduction

Casual empiricism suggests that banks offer two disparate types of deposit liabilities: *transactional*-type demand-deposit account liabilities, and *saving*-type time-deposit account liabilities. Transactional-type accounts provide immediate access to funds, are immediately transferable, pay very little or no interest, charge low or no transaction fees, and provide access to additional credit and overdraft spending facilities. Saving-type accounts offer limited transferability of funds, or are transferable only after some notice period. These accounts offer higher interest rates the longer the notice period, charge transaction fees, and do not provide linked credit facilities. These differences suggest a *prima facie* distinction between accounts representing different underlying intermediation services: accounts that function as a cash substitute used primarily for transacting and the ability to make payments (monetary intermediation), and those that function primarily as savings products of varying maturities (financial intermediation).

In this part of this chapter, I use bank balance sheet data and structural vector autoregression (SVAR) analysis to test whether there are differences in how these liabilities respond to and impact on the broader economy.[22] Results indicate that (i) different account types have different responses to various macroeconomic shocks; (ii)

[22] This study uses South African bank deposit account data available publically from the South African Reserve Bank.

differences in responses correspond to differences we would expect from different underlying intermediation services; and (iii) different account types have different roles in the transmission of monetary policy. These results have at least three policy implications. First, the results indicate that monetary intermediation provides useful insight for explaining the dynamics of bank balance sheets and the transmission of monetary policy. Results suggest that contractionary monetary policy does not constrain the ability of banks to acquire deposits thereby forcing a reduction in bank lending and lowering aggregate expenditure (i.e., a financial intermediation channel), rather they suggest that an increase in interest rates constrains bank credit lending which subsequently lowers expenditure and reduces deposits (i.e., a monetary intermediation channel). This has policy implications. Second, the results imply that regulation based solely on the role of banks as financial intermediaries is worth rethinking. For example, both the US Federal Deposit Insurance Act and the Banks Act of South Africa define the business of banking in terms of accepting deposits and not in terms of *transforming debt into cash equivalent liabilities*. This would be problematic for regulating emerging "fintech" services not premised on the idea of "deposit taking", but instead on providing monetary intermediation services by converting debt into liabilities that are "cryptocurrencies" (as opposed to "deposits"). This would have the potential for exploiting regulatory arbitrage. Third, the results provide empirical evidence of the importance of the role of banks not simply in channelling funds from savers to borrowers, but in providing uniquely monetary transaction services.[23] I am not aware of any other studies that provide similar empirical evidence.

[23] For examples of this literature, see the list of references presented in Lagos, et al. (2017).

The rest of the chapter is structured as follows. Section 3.9 provides a brief look at empirical literature exploring the role of banks in macroeconomic dynamics and policy transmission; Section 3.10 describes the theoretical framework and methodology used; Section 3.11 presents and analyses the results; Section 3.12 discusses some policy implications that follow from this analysis; and finally Section 3.13 concludes.

3.9. Literature

Numerous studies explore the effects bank balance sheets on the transmission of money policy. Bernanke & Blinder (1992) is a seminal study. Using a vector autoregression (VAR) approach, they explore how the size and composition of bank balance sheets in the US respond to changes in the federal funds rate. They find that positive innovations to the policy rate reduce the volume of assets and deposits held by depository institutions. They show that while initially the fall in assets is concentrated in securities, shortly thereafter, these begin to recover, and the decline in deposits becomes reflected almost entirely by a decline in loans. The decline in loans is accompanied by an increase in unemployment.[24] They interpret this as evidence in favour of a bank lending channel of monetary policy transmission.

Kashyap & Stein (1995) use a VAR analysis similar to Bernanke & Blinder (1992), but use disaggregated US bank data to investigate whether capital market imperfections affect the bank lending channel. They find that bank deposits, loans, and securities volumes decrease in response to a rise in the policy rate, but that loans by small banks

[24] Figure 3-A.1 illustrates these findings.

decrease by more than those of large banks. They interpret this as evidence of a bank lending channel driven by a contraction in deposit funding and exacerbated by capital market imperfections.[25]

Drechsler, et al. (2017) explore whether competition in the banking sector affects the transmission of monetary policy in the US. They show that banks in less competitive markets widen the spreads they charge on deposits more than those markets that are more competitive.[26] This increase in spreads then leads to a withdrawal of deposit funding which leads to a decrease in bank lending. They interpret this as evidence in favour of a deposit channel of monetary policy transmission.

Eggertsson, et al. (2019) use a combination of aggregate and bank level data to examine the pass-through of negative interest rates via the banking system in Sweden. They find that negative policy rates have limited pass-through to deposit rates, and that deposit rates are bounded close to zero. They argue that because depositors have the option of holding cash as an alternative store-of-value to deposits, and because cash pays a zero interest rate, therefore deposits rates are constrained by this zero lower bound; i.e., depositors can simply substitute into cash if rates are lowered below zero. They also find that alternative forms of bank funding (for example covered bonds) are not constrained by this zero lower bound and that in some cases bank lending rates increased rather than decreased in response to policy rate cuts. As a result, they find that a lack of pass-through to lending rates can be explained by cross-sectional variation in banks' reliance on deposit financing. They show banks that rely more heavily on

[25] See Figure 3-A.2, which illustrates their key findings. A follow-up study by Kashyap & Stein (2000) further support these results.
[26] Spreads are measured in relation to the risk-free policy rate.

deposit financing are less likely to reduce their lending rates and that this constrains the pass-through of negative policy rates.[27]

The results of Eggertsson et al. (2019) raise two questions. First, why do banks not substitute away from relatively expensive deposit funding and replace it with cheaper non-deposit funding? Eggertsson et al. do not provide a definitive answer, but speculate that it could be because maintaining a base of depositors creates some synergies which other financing sources do not; or that the room for new issuances of covered bonds may be limited by the availability of bank assets to use for collateral; or that Basel III regulation makes deposit financing more attractive in terms of satisfying certain requirements. The second question is why are a bank's non-deposit funding options not equally constrained by a zero lower bound? Surely investors in these alternative funding options could also simply invest in cash to avoid a negative return? The authors do not pose or address this question, but one could speculate from a financial intermediation perspective that there are some costs to investing in cash that outweigh the negative returns on these alternative funding options.

As an alternative framework for answering both of these questions, we can turn to the transaction services that banks provide. In answer to why only deposit funding is subject to a zero lower bound, note that from a transactions perspective, only a bank's deposit liabilities offer a *transaction* service substitute to cash; and that whereas deposits provide a transaction substitute for zero-interest cash, cash is an ill-suited *investment* substitute for other forms of bank funding. That is, cash is not easily traded or transferred in large quantities or across large distances without significant transport

[27] The authors do not report on how these banks' deposit liability types change as a result of the negative interest rates.

costs. These characteristics explain why a zero lower bound constraint only applies to deposit liabilities which provide a substitute to cash. In answer to why banks do not substitute away from relatively expensive deposit funding, note that bank deposit liabilities are unique specifically because of their transaction services. If banks were to substitute away from deposits, it would mean giving up their services as monetary intermediaries. Therefore, despite the zero lower bound, banks are unable to substitute away from deposits for cheaper forms of financing without sacrificing these transaction services. If these alternative forms of financing did provide transaction services, then they would compete directly with cash and would equally be constrained by a zero lower bound, thus losing their cost advantage—a catch-22 situation. The findings of Eggertsson, et al. (2019) therefore suggest that bank transaction services provide a useful framework for understanding some important aspects of bank balance sheet dynamics. But there are no empirical studies (that I am aware of) that focus specifically on how transaction services impact on monetary policy transmission and the economic effects thereof.

3.10. Theoretical framework and methodology

The theoretical framework is based on the following thought experiment: if banks provide both monetary (transaction) and financial (savings) intermediation services, could one expect to see differences in the way that these services affect bank balance sheet dynamics? Consider the following scenarios.

Consider the effects of monetary policy interest rate innovations on deposit account *interest rates* and *balances*. If banks provide monetary intermediation services (such that deposit accounts serve as a medium-of-exchange substitute to cash), one would *not* expect banks to pay interest on these deposits given that cash pays no interest. Therefore, there would be *little reason for changes in the policy rate to impact on rates paid* on these transactional-type deposits (if rates are paid). With regards to account balances, there are two points to consider. First, the transaction demand for money is traditionally thought of as perfectly interest inelastic—especially in the short run and with low market rates (Hansen, 1949; Baumol, 1952; Tobin, 1956). This is standard monetary theory. Second, a rising policy rate is expected to constrain demand for credit-financed expenditure due to higher real interest rates that dissuade borrowing—that is, the “interest rate channel” of monetary policy transmission (Mishkin, 1995; Ireland, 2008). Therefore, if raising the policy rate constrains credit-financed expenditure, one would expect lower demand for monetary intermediation services—that is, less need for converting buyers’ debt into money for sellers. One would therefore *expect lower transactional-type account balances as a result of a higher policy rate*, even if the transaction demand for deposits is interest inelastic (i.e., a shift in the interest inelastic transactions demand curve).

If, instead, deposits are held as a savings instrument, the portfolio theoretic approach of financial intermediation theory predicts that, due to competition from financial markets and other financial intermediaries, banks must pay *market related interest rates* on these saving-type deposit accounts (Fama, 1980). If they do not, these deposit funds will shift to other investment products (an asset demand for deposits). In well-developed and competitive markets, one would therefore *expect saving-type deposit interest rates to closely track changes in the policy rate*. With regards to saving-type deposit balances, there are a number of points to consider. To the extent that contractionary monetary policy withdraws deposit funding from banks and this leads to a contraction in bank lending—that is, the bank deposit, bank lending, and credit channels of monetary policy (Bernanke & Blinder, 1992; Drechsler, et al., 2017)—and to the extent that this reduces expenditure and income growth, one could expect aggregate savings growth, and therefore saving-type account balance growth, to decrease despite the increase in interest rates. This is standard ISLM theory. But in a dynamic setting, the *expected response of saving-type account balances suffers from ambiguity due to the relative impacts of the income, wealth, and intertemporal substitution effects* on the interest elasticity of savings (Giovannini, 1983; Hall, 1988; Campbell & Mankiw, 1989; Elmendorf, 1996; Best, et al., 2018). One might therefore expect the response to monetary policy innovations from both individual banks, and the banking sector more broadly, to differ depending on the proportion of bank balance sheets driven by monetary or financial intermediation services.

Consider next the effects of aggregate demand (AD) innovations in two different cases. First, consider an unspecified negative shock to autonomous spending (AD_1 shock). One would expect this shock to decrease both expenditure and income and therefore to decrease both transactional-type account balances and saving-type account balances (Hansen, 1949; Tobin, 1956). Second, consider a shock that specifically increases the propensity to save (AD_2 shock). One would expect this shock to increase saving-type account balances. However, the paradox of thrift suggests that such a savings shock would also depress general economic activity. It would decrease expenditure and income and therefore decrease both transactional-type and saving-type account balances. In response to both these negative expenditure shocks, bank balance sheet growth is therefore *expected to contract*, but the more banks are funded by transactional-type accounts, the greater this contraction due to the absence of an offsetting savings effect (especially if the shock is due to an increase in the propensity to save). Furthermore, one would expect an inflation targeting central bank to decrease the policy rate in response to falling inflation as a result of falling expenditure. The expected bank account interest rate responses to this monetary policy reaction follow from the previous discussion. Therefore, banks would once again be expected to react differently to the same spending shock and resultant policy intervention, depending on the extent to which these banks provide monetary or financial intermediation services.

A summary of these responses is presented in Table 3-1. As Table 3-1 illustrates, if different bank account types reflect different underlying intermediation services, one would expect account balance responses to monetary policy shocks (Row i) and savings shocks (Row iii) to vary according to the underlying intermediation services provided. One could expect the response of transactional-type account balances to be negative in both cases, but any negative response from saving-type accounts is expected to be softened by counteracting positive effects. By contrast, one could expect that account balance responses to contractionary expenditure shocks (Row ii) will be similar across different account types in terms of sign. One would expect account interest rates to show very little response to any shocks in the case of transactional-type accounts, which is unlike the response that would be expected of saving-type accounts. In order to test these predictions, identifying appropriate shocks and mapping the responses of bank account characteristics to these shocks is an important aspect of this study.[28]

Table 3-1: Theoretical responses of deposit characteristics to contractionary shocks

Acc. type:	**Saving-type**		**Transactional-type**	
Function:	**Financial intermediation**		**Monetary intermediation**	
Acc. char.:	**Int. rate**	**Balance**	**Int. rate**	**Balance**
Shocks:				
(i) MP:	Repo↑ → rate↑	(i) Repo↑ → AD↓ → inc.↓ → sav.↓ → bal.↓ (ii) Repo↑ → rate↑ → sav.↑ → bal.↑	No effect	Repo↑ → AD↓ → exp.↓ → bal.↓
(ii) Exp.:	AD↓ → infl.↓ → Repo↓ → rate↓	AD↓ → inc.↓ → sav.↓ → bal.↓	No effect	AD↓ → exp.↓ → bal.↓
(iii) Sav.:	Sav.↑ → AD↓ → infl.↓ → Repo↓ → rate↓	(i) Sav.↑ → bal.↑ (ii) Sav.↑ → AD↓ → inc.↓ → bal.↓	No effect	Sav.↑ → AD↓ → exp.↓ → bal.↓

3.10.1. Techniques used

The techniques used to explore how account balance and interest rate characteristics differ across different bank deposit account types include: (i) graphing the time series data, (ii) constructing correlation matrices, and (iii) SVAR analysis. An SVAR specification was chosen as a suitable regression modelling tool for the analysis as it provides a framework for dealing with endogeneity across the variables and for

[28] The discussion assumes that the velocity of money remains fairly constant over the time frame of the study. But if it does not, then evidence of the expected relationships would likely be undermined.

identifying appropriate shocks to the variables. Following Kilian (2013), the SVAR model is specified as:[29]

$$Ay_t = C_1 y_{t-1} + \ldots + C_k y_{t-k} + B\varepsilon_t$$

$$\mathrm{E}(\varepsilon_t \varepsilon_t') = \Sigma = \mathrm{I}$$

where y_t is a vector of endogenous variables; A, B, and $C_{i\ (i\,=\,1,\,\ldots,\,k)}$ are coefficient matrices and the parameters of interest in the study; ε_t is a vector of exogenous shocks; and Σ is the covariance matrix of shocks which are restricted to an identity matrix (i.e., the shocks are constructed to be contemporaneously uncorrelated).[30]

The vector of endogenous variables modelled consists of the consumer price index (*Prices*), the policy rate (*Repo*), aggregate expenditure (*Exp*), and bank deposit account balance (*Acc. balance*) and interest rate (*Acc. int. rate*) characteristics. Nominal account balance and expenditure data were transformed to real data by adjusting for the price index. The *Prices*, *Acc. balance*, and *Exp* variables were adjusted for non-stationarity and seasonality by taking the 12-month difference of the natural logarithm. These variables are therefore measured as approximate percentage changes; that is, *Prices* measures inflation, and both *Acc. balance* and *Exp* measure growth rates. Unit root tests confirm the adjusted data are stationary. *Acc. int. rate* and *Repo* are 12-month differenced for consistency. Various information criteria (IC) were calculated to choose an appropriate lag order. These did not provide a clear indication of lag order, and so

[29] No deterministic trend is specified because all variables are differenced form.

[30] A distinguishing feature of the SVAR methodology is that structural innovations are assumed to be uncorrelated. Bernanke (1986, p. 52) justifies this restriction by explaining these shocks as "'primitive' exogenous forces, not directly observed by the econometrician, which buffet the system and cause oscillations. Because these shocks are primitive, i.e., they do not have common causes, it is natural to treat them as approximately uncorrelated".

lag order was chosen according to consensus between the Hannan and Quinn IC, Schwarz's Bayesian IC, and by plotting the autocorrelation and partial autocorrelation functions of all the variables. Based on this analysis, a single lag was chosen as the most suitable.

Solving the SVAR model means estimating A, B, and C_i. However, without additional assumptions, no conclusions about these structural parameters can be drawn from the reduced form estimates, because different structural models will give rise to the same reduced form model (Gottschalk, 2001). A set of restrictions is therefore required to identify the structural parameters. A popular approach for imposing the requisite number of restrictions is the Cholesky identification method, which restricts matrix A to a lower triangular matrix, and B to an identity matrix (Kilian, 2013). This imposes a recursive structural ordering on the system. In terms of this structural ordering, shocks to equations ordered lower in the matrix structure are restricted from having a contemporaneous effect on equations ordered higher in the matrix structure. Shocks to equations ordered higher in the matrix structure are not restricted from having a contemporaneous effect on equations lower in the matrix structure. This is the approach taken in this study. The system of equations is therefore given by the following matrix structure:

$$\begin{bmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \\ * & * & * & * & * \end{bmatrix} \begin{bmatrix} Prices_t \\ Repo_t \\ Acc.\ rate_t \\ Exp_t \\ Acc.\ bal._t \end{bmatrix} = \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \begin{bmatrix} Prices_{t-1} \\ Repo_{t-1} \\ Acc.\ rate_{t-1} \\ Exp_{t-1} \\ Acc.\ bal._{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \varepsilon_t$$

The recursive matrix ordering imposes short-run structural identifying restrictions on the shocks to the system. These restrictions are imposed on *a priori* grounds and must therefore be based on justifiable theoretical foundations (Cooley & LeRoy, 1985).[31] These are justified as follows.

Placing the *Prices* equation above the *Exp* equation imposes the restriction that shocks to *Exp* have a contemporaneous effect on the real value of expenditure (i.e., quantity), but not on the price level. In other words, prices are treated as sticky in the short run. This restriction is justified by noting that the assumption of sticky prices is a standard theoretical underpinning of monetary policy theory. Sticky prices imply that policy changes to nominal interest rates affect real interest rates so that spending decisions adjust accordingly. This restriction means that aggregate supply (AS) is modelled as perfectly price elastic in the short run, so that shocks to *Prices* are identified as shocks, whereas shocks to the *Exp* equation are identified as AD shocks.

Exp is treated as a proxy for AD, and by extension as a proxy for the transactions demand for money (MD_1). The *Exp* equation is placed above the *Acc. balance* equation, imposing the restriction that shocks to *Exp* have a contemporaneous effect on *Acc. balance*, but not *vice versa*. Placing the *Acc. balance* equation below the *Exp* equation is intended to identify innovations to account balances that are orthogonal to transaction demand shocks. The hypothesis is that innovations that are orthogonal to transaction demand shocks will include shocks to the asset demand for money (i.e., savings shocks—MD_2). These orthogonal shocks are expected to be primarily savings shocks in the case of saving-type account balances, whereas they are expected to capture omitted variable effects in the case of transactional-type account balances.

[31] The model remains subject to the otherwise standard SVAR limitations. These include concerns about whether the shocks indeed measure what they are claimed to measure, and whether the proposed shocks should be regarded as orthogonal in the first place (Gottschalk, 2001). These concerns remain important caveats to the analysis.

Ordering the *Repo* equation above both the *Exp* and *Acc. balance* equations (i.e., above the money demand equations) assumes that money demand shocks do not affect the policy rate, implying a perfectly interest-elastic money supply curve. This restriction is rationalised by the fact that the South African Reserve Bank (SARB) is an inflation targeting central bank that directly sets and controls the policy rate. Therefore, shocks to money demand (via the *Exp* and *Acc. Balance* equations) have a contemporaneous effect on the quantity of money, but not on the policy rate. Changes to *Repo* are therefore identified as discretionary changes to monetary policy. Because the *Repo* equation is ordered below the *Prices* equation, but above the *Exp* equation, the assumption is that, while monetary policy decisions are not a factor that contemporaneously influences supply side dynamics, they do impact contemporaneously on spending decisions.

Finally, the *Acc. int. rate* equation is placed above the *Exp* and *Acc. balance* equations, but below the *Repo* equation. Placing it below the *Repo* equation identifies interest rate shocks that are orthogonal to policy changes. Because South Africa is a small open economy, these shocks will likely capture various international financial factors—including global liquidity and other global financial cycle effects. These shocks are therefore likely to be difficult to interpret. These shocks are modelled as contemporaneously impacting on expenditure decisions, but not *vice versa*. The intuition here is that these bank funding costs affect spending decisions (via the bank lending channel), but local spending decisions are unlikely to have a contemporaneous impact in the face of broader global funding conditions. Therefore, spending decisions are restricted from having first round effects on costs of funding. This restriction assumes that these funding cost (FC) shocks are a reflection of the costs of raising financial capital (other than the policy rate). The model system of equations (excluding lag terms) is therefore structured as follows:

$$Prices = \varepsilon^{AS} \quad (3.1)$$

$$Repo = \alpha_1 Prices + \varepsilon^{MP} \quad (3.2)$$

$$Acc.\ int.\ rate = \alpha_2 Prices + \alpha_3 Repo + \varepsilon^{FC} \quad (3.3)$$

$$Exp = \alpha_4 Prices + \alpha_5 Repo + \alpha_6 Acc.\ int.\ rate + \varepsilon^{MD_1} \quad (3.4)$$

$$Acc.\ balance = \alpha_7 Prices + \alpha_8 Repo + \alpha_9 Acc.\ int.\ rate + \alpha_{10} Exp + \varepsilon^{MD_2} \quad (3.5)$$

3.10.2. Data

Public bank deposit account balance and interest rate data were obtained from the SARB's BA930 form. The bank account data are monthly and the period used runs from January 2008 until September 2017 for a total of 115 observations of each series. The start period was chosen according the availability of *all* variables used in the analysis. Publically available aggregate macroeconomic data series were obtained from either the SARB, Statistics South Africa (StatsSA), or the Morningstar money market fund (MMF) database. These data series include measures of market interest rates, measures of aggregate expenditure, and measures of the aggregate price level.

Bank deposit data are categorised by the SARB into three sectors: corporate (C), household (H), and foreign (F). Data are available to the public at an aggregate level for each sector account type. Data for individual banks are not available to the public. Each sector has data for the following account offerings: cheque accounts; call deposits; notice deposits (0–1 month, 1–3 months, and 3–6 months); and fixed deposits (0–1 year, 1–3 years, 3–5 years, and 5+ years). The household sector has an additional account offering called a savings deposit account. Month-end balances and weighted average returns are available for each sector account type. Cheque accounts are the only unambiguously transactional-type accounts offering immediate access to all funds, no or very low transaction fees, and overdraft facilities. Call accounts offer the most liquid of the saving-type accounts, and also provide some transaction services by providing immediate access to a small portion of funds. There is usually a short time delay (normally 24 hours) on the available funds. Call accounts charge higher transaction fees than cheque accounts, and do not provide overdraft facilities. Savings accounts offer a hybrid of call and cheque account features. They offer immediate access to most funds, but charge higher transaction fees than cheque accounts, and do not provide overdraft facilities.

The SARB provides monthly data on the following comparative market interest rates: the repo rate; the South African bank overnight rate (SABOR); 3-month treasury bills; and government bonds (0–3 years, 3–5 years, and 5–10 years). The SARB also provides a measure of the monthly average interest rate charged on bank overdrafts. A monthly average MMF return was calculated for the 66 MMFs available in South Africa. The SARB provides GDP data, but this is only available at a quarterly frequency. While no monthly GDP data are available, monthly expenditure data are available from StatsSA. StatsSA provides monthly retail, wholesale, manufacturing, and mining sales estimates. These were combined to form a measure of aggregate expenditure as a proxy measure for output. StatsSA also provides monthly consumer price index (CPI) measures. The (nominal) data along with various descriptive statistics are presented in Tables 3-2 and 3-3.

Table 3-2: Bank deposit data

		Bank deposit balances (R'm)					Bank deposit interest rates (%)				
	Obs.	Mean	Std. Dev.	Min.	Max.	Source	Mean	Std. Dev.	Min.	Max.	Source
Cheque											
C	115	217.984	81.121	56.754	337.925	8571M	3.194	0.859	2.190	6.530	8573M
H	115	93.548	9.487	42.490	108.278	8681M	1.650	1.473	0.590	5.720	8683M
F	115	10.429	4.801	1.320	42.313	8801M	4.224	1.145	2.960	8.120	8803M
Call											
C	115	286.481	125.483	60.483	487.051	8581M	6.033	1.767	4.350	10.870	8583M
H	115	124.470	44.052	35.340	218.526	8691M	5.580	1.841	3.810	10.580	8693M
F	115	13.061	4.091	2.231	23.131	8811M	5.538	1.983	2.510	10.870	8813M
Savings											
H	115	127.590	30.344	62.987	177.661	8701M	3.597	1.171	2.520	6.780	8703M
0-1m not.											
C	115	37.187	10.477	5.835	54.325	8591M	6.390	1.517	4.830	10.840	8593M
H	115	86.332	25.816	20.145	136.905	8711M	5.601	1.704	3.940	10.000	8713M
F	115	2.317	1.301	0.250	7.283	8821M	5.690	1.664	3.910	10.130	8823M
1-3m not.											
C	115	20.882	9.680	1.386	39.276	8601M	6.979	1.998	5.020	12.420	8603M
H	115	10.724	5.492	0.675	23.652	8721M	6.417	1.989	4.030	11.950	8723M
F	115	0.155	0.274	0.000	2.308	8831M	6.953	2.309	0.000	12.970	8833M
3-6m not.											
C	115	10.079	6.731	0.243	24.704	8611M	6.935	1.309	5.200	10.670	8613M
H	115	3.150	1.646	0.495	6.710	8731M	5.806	1.543	3.990	10.060	8733M
F	115	0.196	0.264	0.000	1.695	8841M	6.214	1.844	2.530	10.510	8843M
0-1y fix.											
C	115	178.402	96.124	23.655	288.735	8621M	7.124	1.891	5.240	11.950	8623M
H	115	83.431	15.832	44.610	114.038	8741M	7.048	1.957	5.130	11.870	8743M
F	115	7.734	3.256	2.478	17.538	8851M	6.804	2.208	2.930	12.010	8853M
1-3y fix.											
C	115	24.585	16.421	0.405	55.733	8631M	8.112	1.681	6.280	12.320	8633M
H	115	20.223	9.430	4.592	46.654	8751M	7.622	1.614	5.840	11.310	8753M
F	115	1.824	1.487	0.050	5.633	8861M	8.058	1.511	6.140	11.560	8863M
3-5y fix.											
C	115	13.265	10.108	0.059	30.874	8641M	8.463	1.224	6.830	11.530	8643M
H	115	6.720	5.082	0.621	16.950	8761M	8.599	0.604	7.730	9.610	8763M
F	115	0.395	0.630	0.005	1.790	8871M	8.113	0.809	6.680	9.590	8873M
5y+ fix.											
C	115	20.179	14.411	0.100	37.640	8651M	8.836	1.383	7.450	12.380	8653M
H	115	5.743	4.626	0.294	16.202	8771M	9.173	0.540	8.510	11.390	8773M
F	115	1.039	0.959	0.008	3.085	8881M	7.641	2.860	1.610	10.110	8883M

Notes: Source refers to SARB form BA930 aggregate BAT data series provided directly by the SARB

Table 3-3: Market related data

	Obs	Market data: Mean	Std. Dev	Min	Max	Source
Int. rates (%)						
Overdrafts	114	10.805	2.677	8.130	17.450	1404M‡
Repo	117	6.729	1.978	5.000	12.000	1401M‡
SABOR	116	6.575	1.976	4.743	11.818	1444W‡
MMF	117	7.552	2.111	5.006	14.107	MS*
3m T-bill	116	6.710	1.704	4.925	11.383	1405W‡
0-3y Gov.	116	7.078	1.568	4.910	12.310	2000M‡
3-5y Gov.	116	7.578	1.172	5.280	11.620	2001M‡
5-10y Gov.	116	7.948	0.776	6.090	10.420	2002M‡
Expenditure (R'b)						
Retail (nom.)	117	60805.6	14246.3	38298.0	109818.0	P6242.1†
Wholesale (nom.)	117	123297.7	27421.3	76095.0	176328.0	P6141.2†
Manuf. (nom.)	117	141259.0	25254.0	93142.0	196522.6	P3041.2†
Mining (nom.)	116	29946.3	5603.5	17263.9	41726.0	P2041†
Prices						
CPI (Dec 2016=100)	117	81.045	12.353	60.200	103.600	P0141†
Infl. (YoY)	115	5.624	1.228	3.400	8.700	7172A‡

Notes:

† Data available from StatsSA: http://www.statssa.gov.za/?page_id=593

‡ Data available from the SARB: https://www.resbank.co.za/Research/Statistics-Home.aspx

* Data calculated using the Morningstar database: http://www.morningstar.co.za/za/

3.11. Empirical results and analysis

3.11.1. Descriptive statistics

Table 3-2 illustrates that corporate and household cheque and call account balances make up the majority share of bank deposit balances. These account balances make up 51% of total deposits (and this proportion increases to 60% when including household savings accounts). Cheque account balances alone account for 23% of total corporate and household deposits. Foreign cheque and call deposits account for only 3% of total cheque and call deposits, and total foreign deposits account for less than 3% of total deposits. Given the small proportion of deposit values consisting of foreign accounts, I limited my focus to the domestic corporate and household sectors for the sake of brevity. Looking across both Tables 3-2 and 3-3 shows that average interest rates paid on corporate cheque accounts are much lower than those paid on call accounts, other short-maturity deposit accounts, and MMFs. For households, the rate is even lower. Average interest rates paid on call accounts are lower than those paid on MMFs.

Figures 3-2 to 3-5 illustrate time series plots of the interest rates offered on the various bank account types and similar market based products. Figure 3-2 plots the interest rates on cheque, call, and household savings accounts along with the repo rate and the average rate charged on overdrafts. Figures 3-3 and 3-4 plot the repo rate along with the interest rates on bank notice and fixed deposits accounts for corporates and households respectively. Figure 3-5 plots the repo rate and the various short- and long-maturity market rates. Figures 3-2 to 3-5 and Tables 3-4 and 3-5 illustrate that market interest rates, bank saving-type account rates, and bank rates charged on overdrafts all closely track the repo rate. The shorter the notice period of the saving-type account, the higher the correlation coefficient relative to the repo rate—except for cheque accounts. The returns on all saving-type accounts are highly correlated with one another, whereas the correlation coefficients between cheque account rates and the various short-maturity market rates are 10 to 15 percentage points lower than other short-maturity rates. Given the high levels of correlation between the different saving-type accounts, the rest of this comparative analysis focuses only on cheque accounts, call accounts, and household savings accounts.

Figure 3-2: Short-maturity bank deposit interest rates

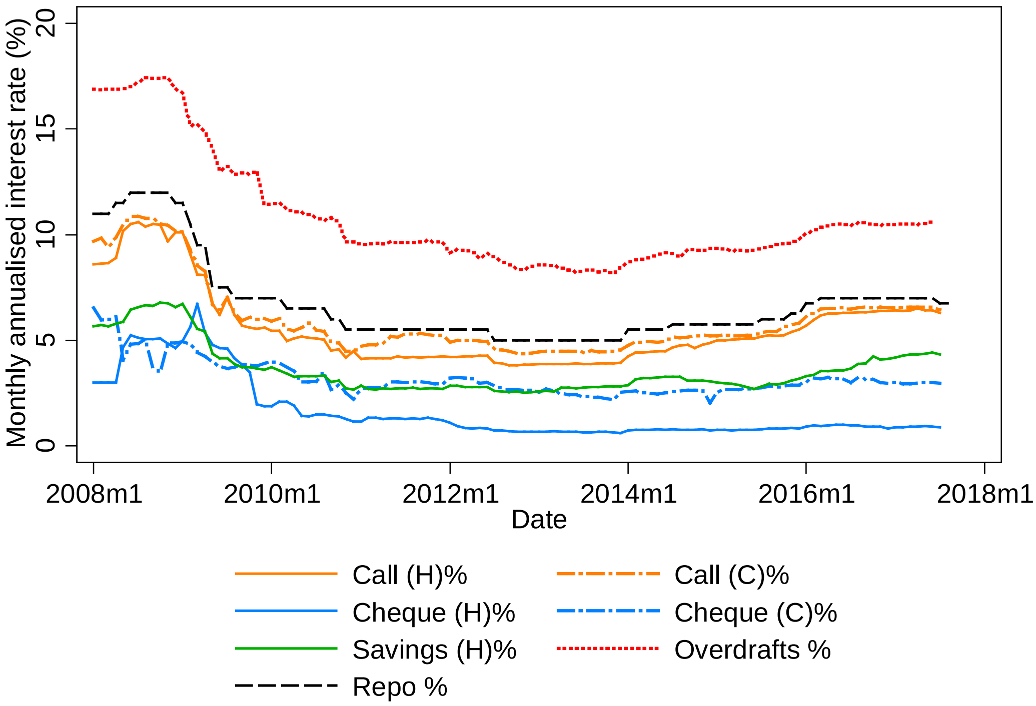


Figure 3-3: Long-maturity bank deposit interest rates (corporate)

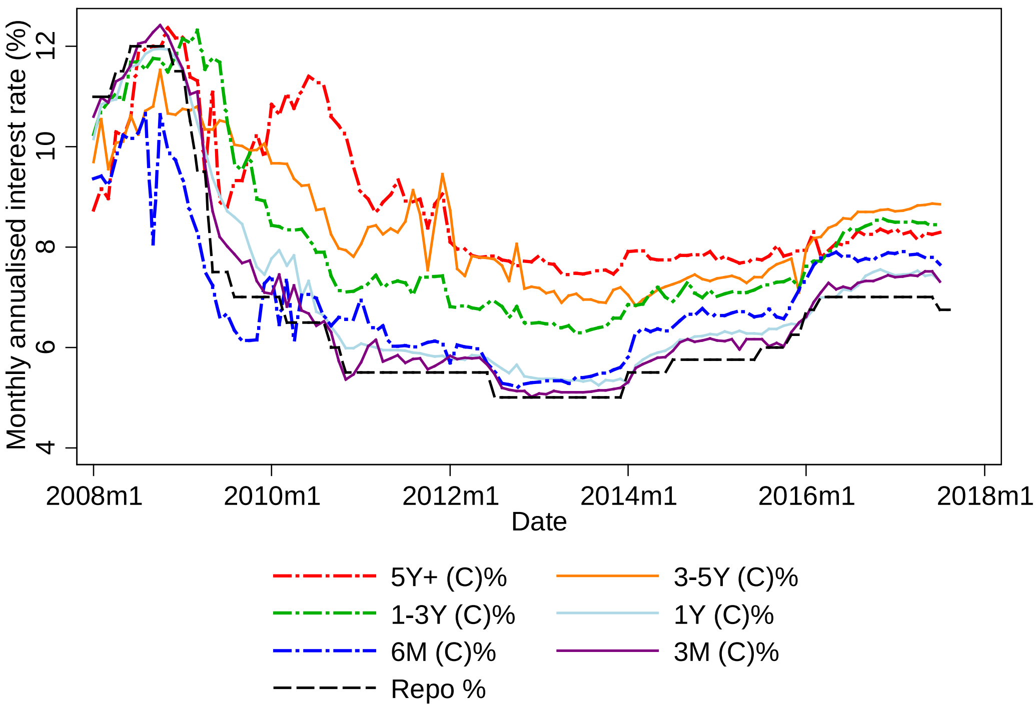


Figure 3-4: Long-maturity bank deposit interest rates (households)

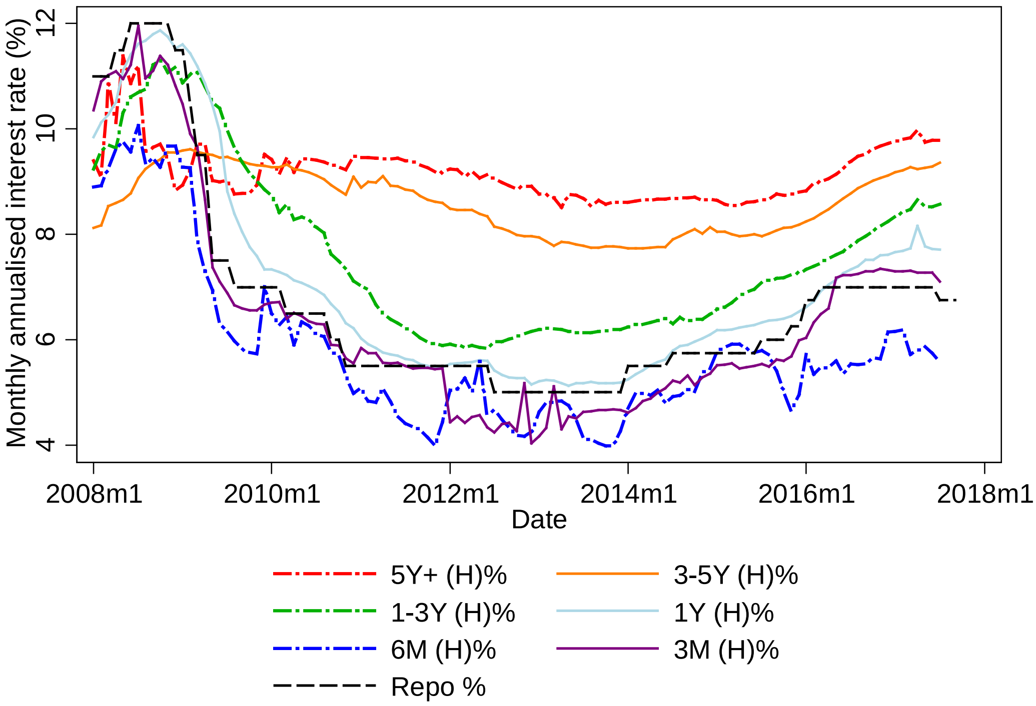


Figure 3-5: Market interest rates

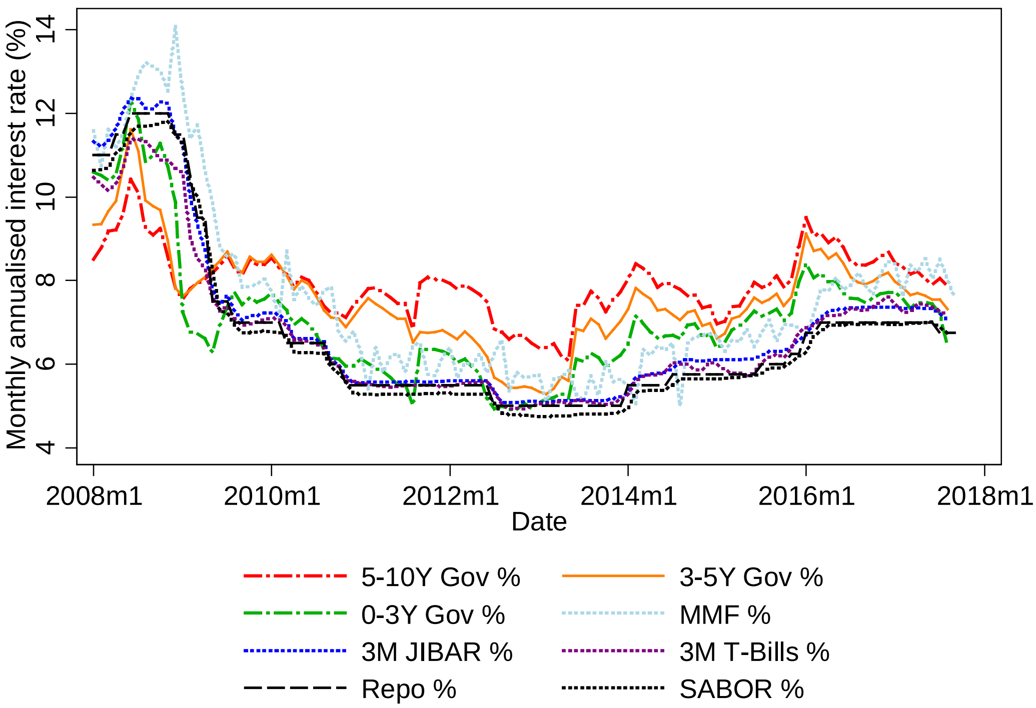


Table 3-4: Short-maturity interest rate correlation coefficients

Variable (%)		(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)	(XIII)	(XIV)
Repo	(I)	1.000													
SABOR	(II)	0.997	1.000												
3m JIBAR	(III)	0.996	0.995	1.000											
3m T-bill	(IV)	0.989	0.990	0.994	1.000										
MMF	(V)	0.969	0.978	0.966	0.967	1.000									
Cheque C	(VI)	0.864	0.858	0.853	0.843	0.825	1.000								
Call C	(VII)	0.993	0.992	0.992	0.989	0.962	0.843	1.000							
0-1m C	(VIII)	0.978	0.984	0.977	0.978	0.960	0.842	0.984	1.000						
1-3m C	(IX)	0.991	0.995	0.987	0.982	0.977	0.855	0.986	0.981	1.000					
Cheque H	(X)	0.843	0.847	0.815	0.798	0.854	0.761	0.820	0.818	0.871	1.000				
Call H	(XI)	0.982	0.985	0.984	0.987	0.967	0.794	0.989	0.984	0.981	0.815	1.000			
Savings H	(XII)	0.977	0.981	0.973	0.975	0.965	0.806	0.976	0.972	0.980	0.831	0.978	1.000		
0-1m H	(XIII)	0.973	0.980	0.975	0.980	0.958	0.799	0.983	0.984	0.973	0.776	0.992	0.977	1.000	
1-3m H	(XIV)	0.981	0.983	0.982	0.981	0.963	0.842	0.978	0.972	0.978	0.807	0.972	0.964	0.965	1.000

Table 3-5: Long-maturity interest rate correlation coefficients

Variable (%)		(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Repo	(I)	1.000											
0-3y Gov.	(II)	0.876	1.000										
3-5y Gov.	(III)	0.798	0.915	1.000									
5-10y Gov.	(IV)	0.664	0.845	0.935	1.000								
0-1y fix. C	(V)	0.981	0.833	0.787	0.638	1.000							
1-3y fix. C	(VI)	0.924	0.731	0.749	0.612	0.969	1.000						
3-5y fix. C	(VII)	0.821	0.634	0.691	0.567	0.883	0.920	1.000					
5y+ fix. C	(VIII)	0.696	0.499	0.543	0.366	0.739	0.740	0.808	1.000				
0-1y fix. H	(IX)	0.963	0.795	0.772	0.625	0.988	0.980	0.881	0.732	1.000			
1-3y fix. H	(X)	0.894	0.739	0.764	0.601	0.948	0.964	0.902	0.765	0.969	1.000		
3-5y fix. H	(XI)	0.529	0.334	0.471	0.386	0.622	0.703	0.838	0.750	0.660	0.743	1.000	
5y+ fix. H	(XII)	0.605	0.555	0.598	0.556	0.584	0.556	0.593	0.486	0.586	0.543	0.576	1.000

Figure 3-6 plots interest rate differentials between: (i) the policy rate and rates paid on call accounts; (ii) the policy rate and rates paid on cheque accounts; and (iii) the rates paid on call accounts and cheque accounts. Figure 3-7 plots deposit balances. Figure 3-8 plots balance differentials between call and cheque accounts (call balance less cheque balance). *Two features stand out.*

First, towards the end of 2011, corporate and household cheque account balances decreased and call account balances increased, seemingly substituting one for the other (Figures 3-7). There was no change in the repo rate at this time (from Figure 3-2). For corporates, this was not a case of funds moving into accounts with higher interest rates as the rate on call accounts relative to cheque accounts fell at this time (Figure 3-6).[32]

Second, from 2012, the return differential between the repo and cheque account rates for both corporates and households has been rising steadily as the policy rate has been rising. By contrast, the rate differential between the repo and call account rates has remained relatively unchanged for corporates, while for households it has fallen. These movements suggest that the underlying determinants for holding balances in these different account types are likely to differ.

[32] The SARB has indicated that this movement was not due to specific regulatory changes, and that it is driven by changes at one major bank. Inserting a dummy into the regression to capture some potential structural change at this point does not materially influence the results.

Figure 3-6: Bank deposit interest rate differentials

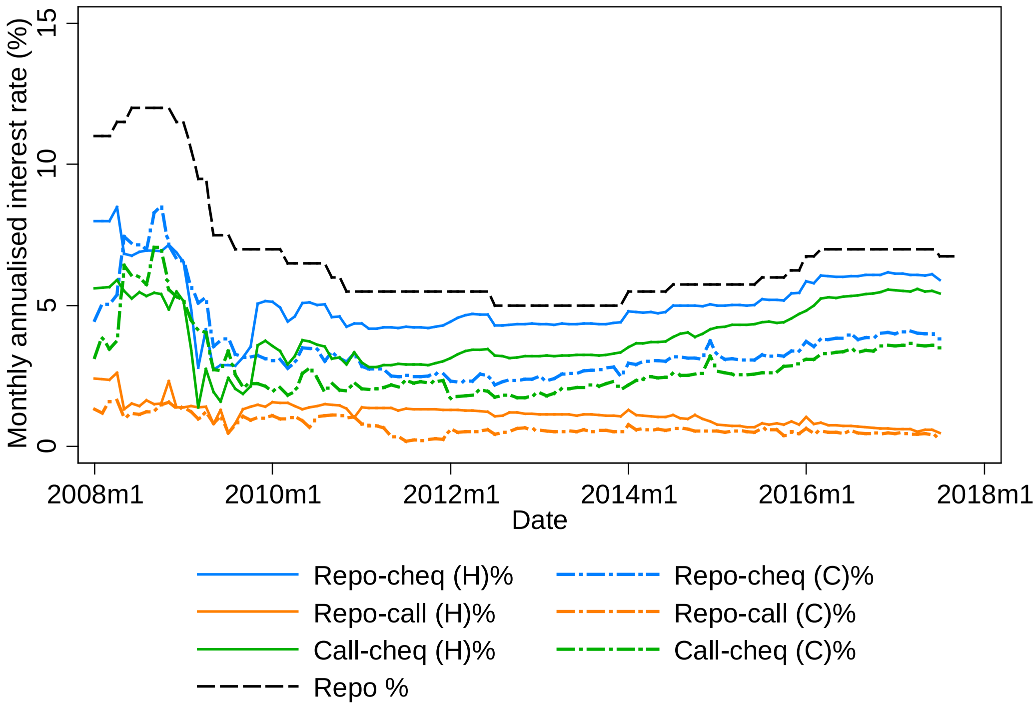


Figure 3-7: Bank deposit balances

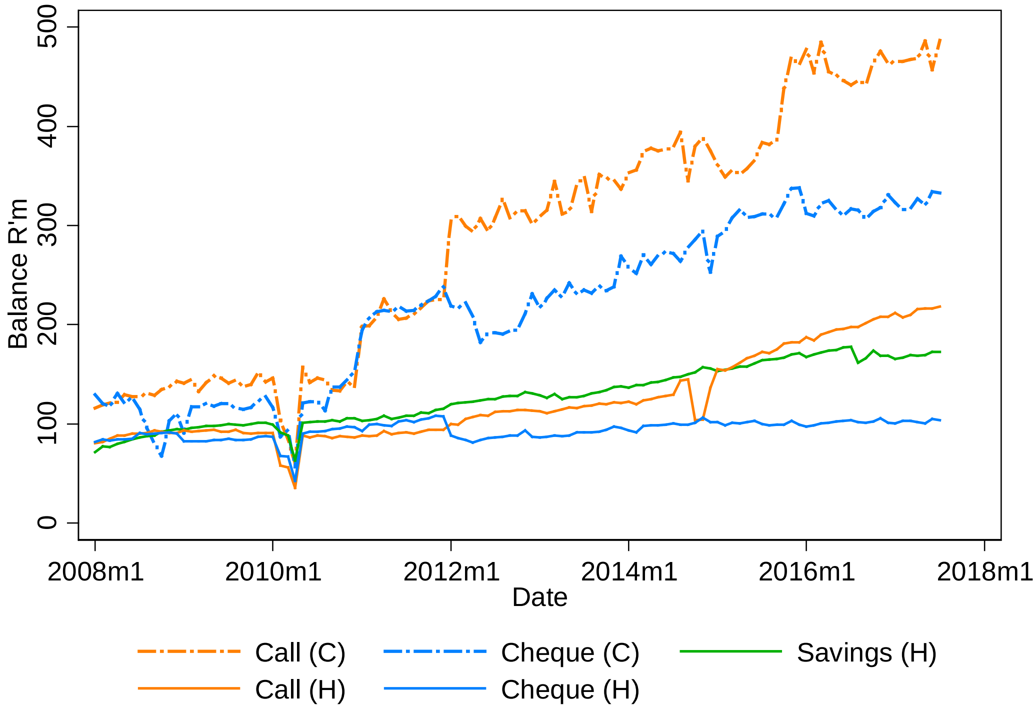


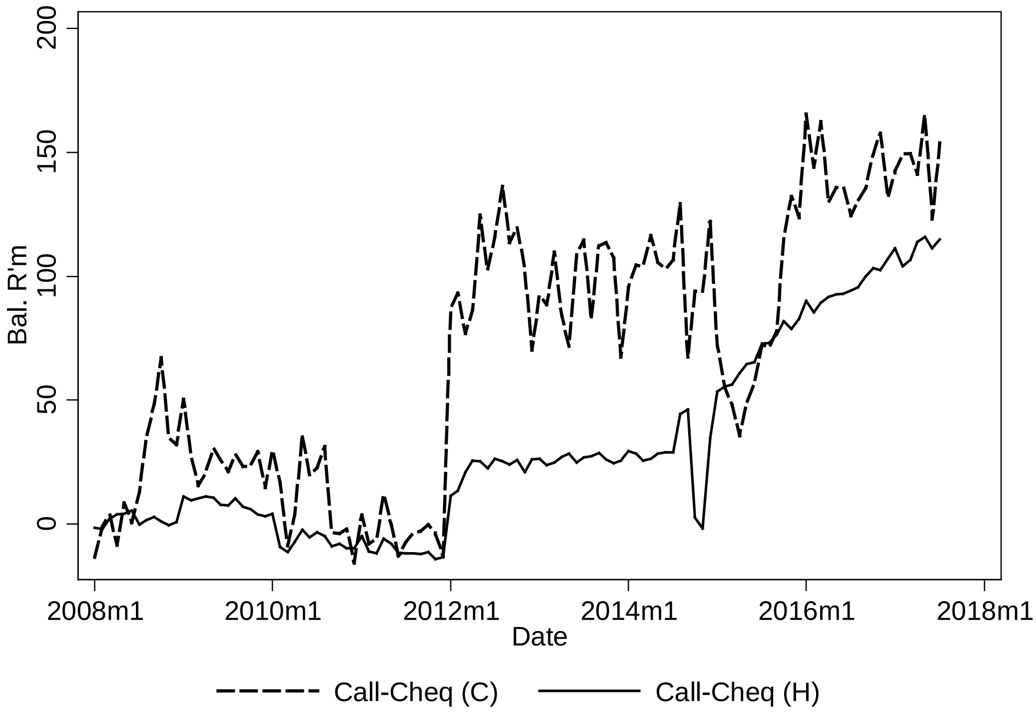
Table 3-6: Correlation coefficients of corporate deposit characteristics

Variable		(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)	(XIII)	(XIV)	(XV)	(XVI)
Repo (%)	(I)	1.000															
MMF (%)	(II)	0.941	1.000														
Overdrafts (%)	(III)	0.935	0.918	1.000													
Cheq. int. rate (%)	(IV)	0.515	0.423	0.398	1.000												
Call int. rate (%)	(V)	0.981	0.916	0.877	0.565	1.000											
Int. rate diff. (%)	(VI)	0.682	0.689	0.662	-0.249	0.659	1.000										
Repo - Cheq. (%)	(VII)	0.738	0.743	0.754	-0.199	0.676	0.975	1.000									
Repo - Call (%)	(VIII)	0.602	0.595	0.741	0.065	0.434	0.450	0.637	1.000								
Cheq. bal. (R'm)	(IX)	-0.406	-0.463	-0.630	-0.019	-0.265	-0.293	-0.449	-0.789	1.000							
Call bal. (R'm)	(X)	-0.466	-0.497	-0.700	-0.225	-0.347	-0.203	-0.356	-0.730	0.885	1.000						
Bal. diff. (R'm)	(XI)	-0.401	-0.393	-0.575	-0.397	-0.342	-0.039	-0.145	-0.451	0.495	0.843	1.000					
Ret. (R'b)	(XII)	-0.197	-0.147	-0.305	-0.116	-0.144	-0.063	-0.133	-0.319	0.430	0.412	0.270	1.000				
Wh. (R'b)	(XIII)	-0.140	-0.194	-0.339	-0.217	-0.063	0.124	0.011	-0.386	0.567	0.641	0.541	0.437	1.000			
Man. (R'b)	(XIV)	0.300	0.230	0.241	-0.081	0.309	0.436	0.406	0.121	0.011	0.045	0.071	-0.014	0.704	1.000		
Min. (R'b)	(XV)	0.166	0.069	0.024	0.271	0.240	0.035	-0.025	-0.219	0.205	0.102	-0.047	0.178	0.509	0.554	1.000	
CPI	(XVI)	-0.410	-0.385	-0.644	-0.221	-0.292	-0.141	-0.295	-0.699	0.829	0.903	0.727	0.420	0.538	-0.010	-0.005	1.000

Table 3-7: Correlation coefficients of household deposit characteristics

Variable		(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)	(XIII)	(XIV)	(XV)	(XVI)
Repo (%)	(I)	1.000															
MMF (%)	(II)	0.941	1.000														
Overdrafts (%)	(III)	0.935	0.918	1.000													
Cheq. int. rate (%)	(IV)	0.683	0.686	0.769	1.000												
Call int. rate (%)	(V)	0.967	0.935	0.866	0.596	1.000											
Int. rate diff. (%)	(VI)	0.558	0.516	0.351	-0.186	0.679	1.000										
Repo - Cheq. (%)	(VII)	0.755	0.671	0.587	0.036	0.787	0.930	1.000									
Repo - Call (%)	(VIII)	0.558	0.444	0.668	0.588	0.327	-0.138	0.236	1.000								
Cheq. bal. (R'm)	(IX)	0.499	0.458	0.621	0.696	0.383	-0.168	0.057	0.600	1.000							
Call bal. (R'm)	(X)	-0.053	-0.027	-0.230	-0.529	0.147	0.663	0.403	-0.669	-0.418	1.000						
Bal. diff. (R'm)	(XI)	-0.222	-0.188	-0.411	-0.671	-0.023	0.587	0.299	-0.748	-0.694	0.944	1.000					
Ret. (R'b)	(XII)	-0.197	-0.147	-0.305	-0.378	-0.113	0.208	0.070	-0.361	-0.176	0.334	0.328	1.000				
Wh. (R'b)	(XIII)	-0.140	-0.194	-0.339	-0.505	-0.061	0.388	0.262	-0.320	-0.236	0.438	0.432	0.437	1.000			
Man. (R'b)	(XIV)	0.300	0.230	0.241	0.018	0.292	0.340	0.393	0.161	0.217	0.085	-0.011	-0.014	0.704	1.000		
Min. (R'b)	(XV)	0.166	0.069	0.024	0.152	0.149	0.043	0.090	0.128	0.240	-0.117	-0.180	0.178	0.509	0.554	1.000	
CPI	(XVI)	-0.410	-0.385	-0.644	-0.740	-0.215	0.414	0.103	-0.818	-0.702	0.801	0.890	0.420	0.538	-0.010	-0.005	1.000

Figure 3-8: Bank deposit balance differentials



Tables 3-6 and 3-7 contain the correlation coefficients of corporate and household deposit account interest rate and balance characteristics, market rates, aggregate expenditure measures, and aggregate price level. Comparing Tables 3-6 and 3-7 illustrates that the correlation coefficients vary not only across account types, but also across firms and households. Table 3-6 shows that both corporate cheque and call account balances are significantly and highly positively correlated with aggregate price levels, positively and significantly correlated with retail and wholesale expenditure, and significantly inversely correlated with interest rates. In Table 3-7, by contrast, household cheque account balances are significantly *inversely* correlated with prices and both retail and wholesale expenditure, and *positively* and significantly correlated with interest rates. Balances across corporate cheque and call accounts are highly positively and significantly correlated, whereas balances across household cheque and call account types are *negatively* and significantly correlated. There is a significant negative correlation between corporate account balances and interest rates, whereas there is significant *positive* correlation between household cheque account balances and insignificant correlation with call account balances. These results suggest that responses to various underlying macroeconomic impulses likely differ across household and corporate cheque and call accounts.

3.11.2. Regression results

SVAR analysis was conducted on all household and corporate deposit accounts. The results by and large support the underlying structural assumptions of the model. Full impulse-response function (IRF) and forecast-error variance decomposition (FEVD) results are presented in Appendices 3B and 3C. Robustness test results are presented in Appendix 3D. The results also support the intuition that the different bank deposit types correspond to different underlying intermediation services. However, results differ across corporate and household accounts, indicating that there are differences in the way households and corporates respond to and transmit shocks. The levels of significance of the results tend to be low, and results across account types are more clearly distinguished for corporates than for households.

3.11.2.1. *Support for the underlying structural assumptions*

IRF results (in aggregate) support the structural identifying restriction that impulses to the *Prices* equation represent AS shocks: positive shocks to inflation are followed by a significant fall in expenditure growth. This is followed by a fall in the policy rate. FEVD results suggest that these impulses explain large and significant portions of the variance of both expenditure growth and the policy rate.

IRF results support the structural identifying restriction that impulses to the *Exp* equation represent AD shocks: positive shocks are followed by a significant rise in the policy rate with a slight rise in inflation that is not significant.

IRF results for the *Repo* equation indicate that a rise in the policy rate is associated initially with a rise in expenditure growth, followed by a rise in inflation. This is *prima facie* counterintuitive, but this so-called "price puzzle" is common in VAR models. Sims (1992) argues that this is a result of specifying VAR models that do not include forward looking variables and therefore do not measure the forward looking component of monetary policy decisions. This price puzzle may therefore capture the contemporaneous correlation between the policy and the response variables (i.e., nominal interest rates rise in environments where expenditure and inflation are rising). The FEVD results provide some support for this interpretation.

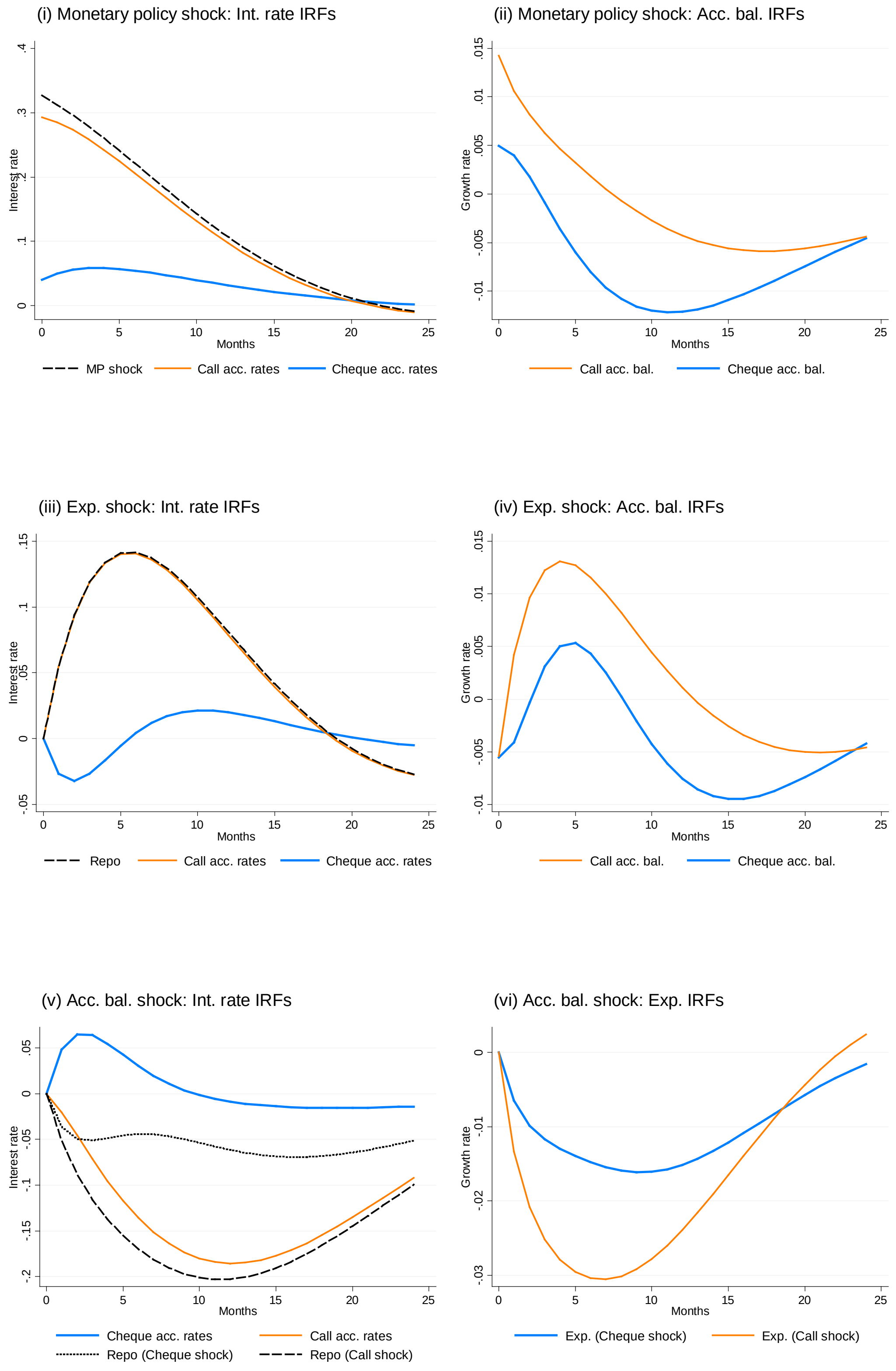
3.11.2.2. *Corporate deposit accounts*

Figure 3-9.i illustrates that both corporate call and cheque account rates exhibit positive significant responses to a positive monetary policy innovation.[33] But whereas the call rate response closely matches the monetary policy impulse, the cheque rate response is significantly less. FEVD plots indicate that monetary policy variance is a significant determinant of a large proportion of call account interest rate variance in the short run, but not so for cheque account rates (see Appendix 3B).[34] Figure 3-9.ii illustrates that cheque account balances have a significant lagged negative response, while call account balances exhibit a smaller negative response that is not significant.[35] These results correspond to those summarised in Table 3-1, Row (i). This suggests that transactional-type accounts and saving-type accounts respond differently to monetary policy shocks, and in a manner that is consistent with the intuition that these different deposit types reflect monetary and financial intermediation services respectively.

A positive expenditure shock is followed by a significant rise in the policy rate, and this is matched closely by a significant call account rate response (Figure 3-9.iii).[36] This is not the case for the cheque rate response, which is negative but not significant. Account balances have a positive response but they are not significant (Figure 3-9.iv),[37]

[33] Confidence intervals are illustrated in Appendix 3B. Confidence intervals are estimated using Stata's built in impulse response function estimation command.
[34] Monetary policy innovations remain a large and significant determinant of long-maturity saving-type accounts, but the time taken for a policy shock to explain large and significant proportions of saving-type account rates increases as the period to maturity increases. This pattern is consistent with, and likely explained by, the fact that pre-shock fixed rates paid on previously issued long-maturity deposits will weigh on average rates paid on long-maturity accounts.
[35] There is evidence of a significant negative response for long-maturity savings-type accounts, but the evidence is not consistent across different maturities.
[36] The magnitude of the account rate response for saving-type accounts tends to decrease as maturity increases.
[37] Cheque account balances do, however, respond positively and significantly to retail expenditure shocks (see Figure 3-D.1.i). And there is evidence of a positive significant response in long-maturity accounts.

Figure 3-9: Corporate deposit account summary IRFs



and expenditure shock variance does not explain significant variance in account balances (see Appendix 3B). Because the response of account balances is similar, but expected to be more significantly positive, these results only partially support those summarised in Table 3-1, Row (ii), and so are partially consistent with the intuition that these different deposit types reflect different underlying intermediation services.

Positive shocks to short-maturity saving-type account balances have the characteristics of a positive savings shock: there is a significantly negative expenditure response along with a significantly accommodative policy response that is closely matched by a negative account rate response (inflation is initially positive and becomes negative despite the falling interest rate).[38] These responses are all consistent with a savings shock. By contrast, a positive shock to cheque balances does not elicit these responses. This suggests that these different account balances capture different underlying shocks.

FEVD plots (Appendix 3B) indicate that cheque account interest rate variance accounts for a significant and increasingly substantial proportion of *Repo* variance. The implied direction of causality here is clearly problematic, and suggests that shocks to cheque account rates reflect some omitted factor resulting in induced correlation between cheque rates and *Repo*. Given the structure of the model, this could indicate that the policy rate variance in South Africa is explained by factors other than just domestic inflation and expenditure variance. For example, global monetary policy conditions—the so-called policy dilemma or "irreconcilable duo" (Rey, 2018)—or

[38] The responses are not the same with respect to shocks to long-maturity saving-type accounts.

reactions to other global financial cycle conditions (Algu & Creamer, 2017). But more research is required to pin down these additional factors.

These results suggest that the innovations to short-maturity saving-type account balances consist to a large extent of domestic savings shocks, but this is not so for cheque accounts. The model structure does not allow a direct test of whether a savings shock has a negative impact on cheque account balances, but the results do indicate that a negative expenditure shock likely leads to a decrease in cheque account balances. This would support the results in Table 3-1, Row (iii).

The SVAR results therefore indicate that different corporate deposit account types have different responses to shocks in the broader macroeconomy in South Africa.

3.11.2.3. Household deposit accounts

Figure 3-10.i illustrates that household call, cheque, and savings account rates exhibit significant positive responses to a positive monetary policy innovation.[39] The call rate response closely matches the monetary policy impulse, whereas the cheque rates have a lagged response of much lower magnitude. The response of savings rates lies between that of call and cheque accounts. FEVD plots indicate that monetary policy variance is a significant determinant of a large proportion of saving-type account rate variance, especially in the short run, but monetary policy variance only explains significant

[39] Confidence intervals are illustrated in Appendix 3C. Confidence intervals are estimated automatically using the built in Stata impulse response function estimation command.

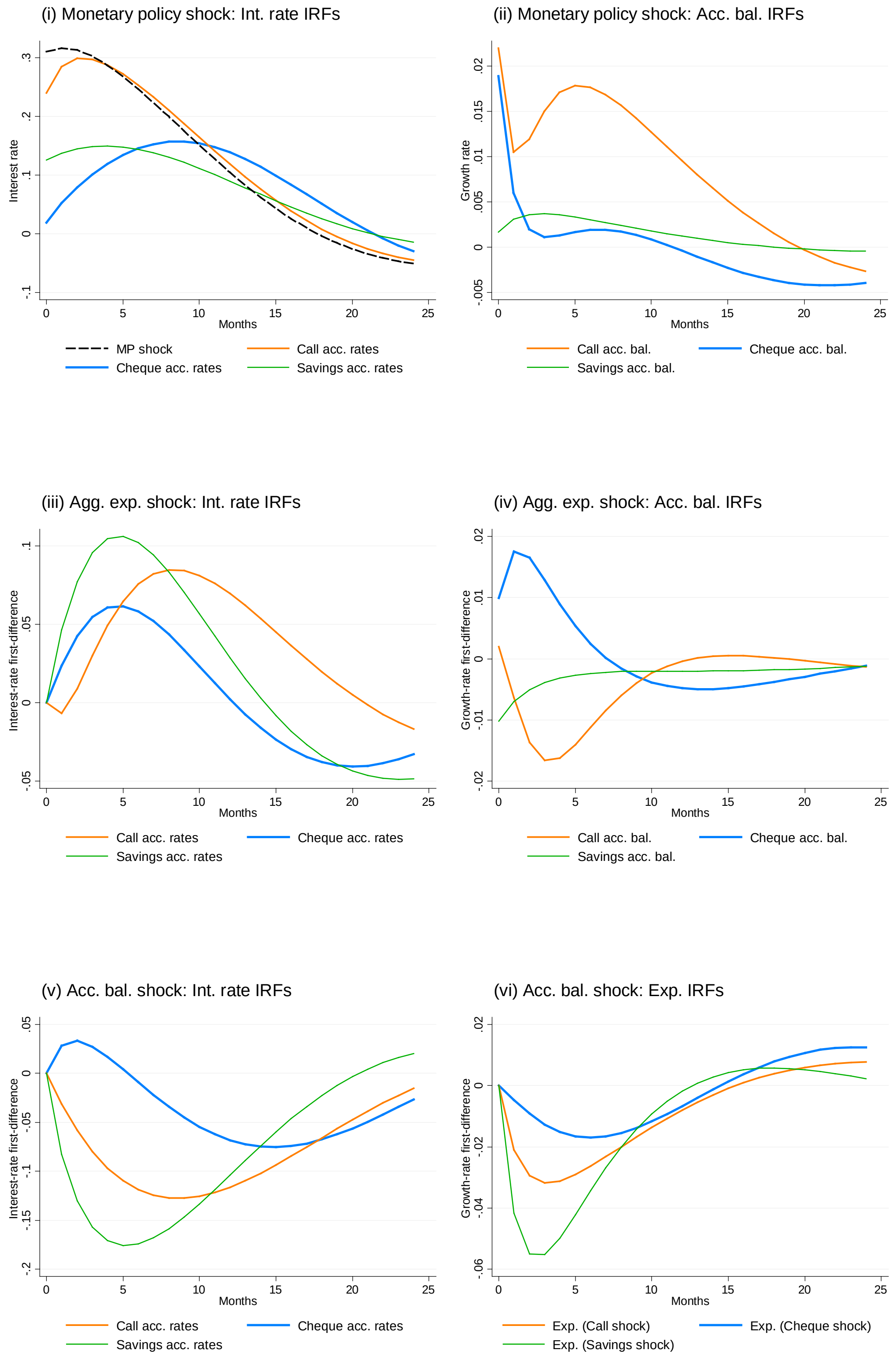
cheque account rate variance after a lag of about a quarter (see Appendix 3C).[40] Unlike cheque and savings balances, call account balances exhibit a positive and significant response to monetary policy innovations.[41] This result is counter to findings using US data (see the discussion in Section 3.9). A potential interpretation is that this reflects the carry trade effects of foreign purchases of South African securities in response to a rise in South African interest rates (Hassan, 2015),[42] but further investigation is required to accurately interpret this result. Monetary policy variance is a significant determinant of call account variance, but not for cheque or savings account balance variance (see Appendix 3C). These results suggest that, like corporate accounts, household transactional-type accounts and saving-type accounts respond differently to monetary policy shocks. The results partly correspond with the anticipated results summarised in Table 3-1, Row (i).

An AD shock (Figures 3-10.iii and .iv) is followed immediately by a rising savings account interest rate response, a smaller and lagged rising call account interest rate response, and an even lower and rising cheque account interest rate response. Only the savings account interest rate response is significant. Both savings and call account balances fall in response to a positive AD shock (but these responses are not significant), whereas cheque balances exhibit a significant rise.[43] The AD shock variance

---

[40] Like corporate accounts, monetary policy innovations remain a large and significant determinant of long-maturity saving-type accounts, but the time taken for a policy shock to explain large proportions of saving-type account rates increases as the period to maturity increases.
[41] Like corporate balances, there is evidence of a significant negative response for long-maturity accounts.
[42] In other words, the lack of negative account balance responses may be because South Africa is a small open economy that attracts significant international financial flows when interest rate differentials rise.
[43] In response to a retail expenditure shock, all account balances rise, although only household savings accounts rise significantly.

Figure 3-10: Household deposit account summary IRFs



does not explain any significant variance in these balances (see Appendix 3C). The cheque account results correspond to Table 3-1, Row (ii)—call and savings account results less so.

Like corporate accounts, positive shocks to short-maturity household saving-type account balances have the anticipated characteristics of positive savings shocks (Figures 3-10.v and .vi). The call and savings account results correspond to the anticipated results summarised in Table 3-1, Row (iii), whereas cheque account balance shocks are less clearly interpreted.

3.11.2.4. Monetary policy transmission

The discussion in Sections 3.11.2.1 to 3.11.2.3 suggests that monetary policy decisions in South Africa are forward looking. They appear to have been effective at mitigating demand side inflation.[44] Monetary policy also appears to have responded to the negative growth effects of supply side shocks.[45] All three of these results provide possible explanations for why monetary policy variance does not explain significant inflation variance in the model, despite this being an anticipated outcome of inflation targeting policy.

Positive monetary policy innovations coincide with a positive contemporaneous expenditure response, reflecting an output puzzle that is common to the VAR literature. They coincide with a contemporaneous positive and significant response from

[44] An alternative explanation is that there has simply been very little demand side inflation during this period.
[45] This result may be driven by the negative real effects of the financial crisis during the period.

household call account balances that may reflect carry trade effects, but they also coincide with contemporaneous negative and significant responses from long-maturity household and corporate saving-type account balances. Corporate cheque account balances have a negative and significant *lagged* impulse response, and this corresponds with a negative lagged expenditure response.

These results suggest that different deposit types have different roles in monetary policy transmission, and that transmission differs across the household and corporate sectors. In response to contractionary monetary policy, long-maturity saving-type accounts contract.[46] But this reduction in deposit funding does not coincide with, nor is it closely followed by, a contraction in spending. However, after a lag of between 1 and 3 quarters, corporate cheque account deposit balances contract, and this does coincide more closely with a contraction in expenditure. These cheque account and expenditure contractions do not occur contemporaneously with a rise in the policy rate, nor contemporaneously with cheque account interest rate spread increases. It therefore seems *unlikely* that these contractions reflect either a bank lending channel or a deposit channel of transmission (i.e., a financial intermediation channel); especially given that household call account balances rise significantly in response to a rise in the policy rate. In other words, it seems *unlikely* that the decrease in transactional-type deposit balances is a reflection of a withdrawal of funding. It seems more likely that the increase in interest rates constrains demand for credit financed expenditure, and that this lowers demand for monetary intermediation services, which contracts both transactional-type

[46] Corporate accounts make up nearly 70% of the value of account balances that have maturities of more than one month; see Table 3-2. The SARB does not conduct open market operations when implementing monetary policy decisions, so it does not seem likely that this captures funds moving off bank balance sheets and into these government securities.

account balances and aggregate economic activity simultaneously. One can think of this as a monetary intermediation channel of policy transmission.[47]

3.11.2.5. Robustness tests

To test the robustness of the results, I experimented with various alternative specifications. I tried a number of different recursive orderings, including completely reversing the ordering of the model and therefore imposing restrictions opposite to those reasoned as plausible (and the economic interpretations of which seem theoretically unjustified). Apart from changes to initial responses (by construction), there were no dramatic differences to the shapes of the IRFs thereafter (see Figures 3-D.1.v and 3-D.1.vi as examples). I compared results of the model with one lag period, three lag periods, six lag periods, and 12 lag periods. The results with only one lag were much smoother. Additional lags appeared to add little more than noise (see Figures D.1.vii and D.1.viii as examples). To better understand the lack of significant response of *Acc. balance* to aggregate AD shocks, I split aggregate expenditure into its subcomponents of industry expenditure. Account balance responses to wholesale, manufacturing, and mining expenditure impulses gave similar results to the aggregate results. But the account balance response to retail sales shocks differed considerably. Account balances respond positively and significantly to retail expenditure shocks, but the variance explained by the shocks is not significant—see Figures 3-D.1.i to 3-D.1.iv

[47] The model setup does not allow one to rule it out as an exchange rate, asset, or wealth channel, but regardless, the results indicate that the transaction services of banks provide useful insight into the dynamics of bank balance sheets and the transmission of monetary policy.

and Figure 3-D.2. I swapped the *Repo* variable for a proxy by using SABOR interest rates (Figure 3-D.1.ix), and again the results did not noticeably differ.

3.12. Policy discussion

The results presented provide support for the intuition that banks provide both monetary and financial intermediation services. Evidence of the effects of a bank's unique monetary intermediation role provides a rationale for treating banks differently to other financial intermediaries, and for developing policy and regulation targeted specifically at this monetary function. This has implications for both monetary and financial stability policy. For example, "Regulation Q" of the US Bank Act restricted the interest rate paid on demand deposits. If, hypothetically, this restriction were taken further and banks were restricted to offering *only* zero-interest on-demand-and-at-par deposit accounts, this would in effect prevent banks from providing financial intermediation services, and restrict banks to providing only monetary intermediation services. Such regulation would have at least two interesting implications related to issues of monetary and financial stability. The first relates to the response of bank balance sheets to policy-induced changes in interest rates. The second relates to efforts central banks may make to change the shape of the yield curve (through, for example, the Fed's Operation Twist). The following discussion illustrates these points.

The response to interest rates: An increase in interest rates that decreases demand for bank loans, thereby shrinking bank balance sheets and putting pressure on loan affordability, would simultaneously increase bank lending margins (given that banks pay no interest on deposits in this scenario). These counteracting effects would minimise the negative impact on bank profits. Contractionary monetary policy would therefore have a limited negative impact on the stability of the banking sector relative to the case where banks pay interest on deposits. When banks pay interest on deposits, an increase in interest rates is expected to shrink growth in bank loan assets, and also negatively impacts profitability by reducing funding margins. This negatively impacts on banking sector stability.[48] If banks were restricted to zero-interest accounts, rising bank margins would offset shrinking bank balance sheets. This suggests that stability could be supported by regulation aimed at separating a bank's monetary and financial intermediation activities (as was the case in the 1930s). But an important caveat is the fact that banks are specifically restricted from paying interest on deposits in this case. It is easy to envisage (and based on history, perhaps foolish not to expect) that competition across banks and other financial intermediaries would prompt banks to try and circumvent this regulation by offering *de facto* interest earning deposits through other implicit-interest reward systems framed, for example, as loyalty benefit programmes. It is not clear how this sort of creativity could be effectively regulated.

The effect on maturity transformation: Stein (2013) describes how maturity transformation can play a role in propagating systemic instability, noting that "it is not just bad credit decisions that create systemic problems, but bad credit decisions combined with excessive maturity transformation" (Stein, 2013, p. 5). Of concern here is the potential for instability when risky long-maturity credit is financed by short-maturity demandable claims that can be recalled if things start to "look bad" (Stein,

[48] The case of the Savings and Loan crisis in the US during the 1980s provides an illustrative example of the potential severity of this problem.

2013). To discourage banks from this "reaching for yield" behaviour, central banks can enter markets to sell short-maturity treasuries and buy long-maturity treasuries with the proceeds, thereby decreasing the spread between short-maturity and long-maturity rates, flattening the yield curve, and sterilising the bond transactions. This would reduce the appeal of maturity transformation by putting a squeeze on margins earned from funding long-maturity credit with short-maturity financing (Stein, 2013). This policy intervention would, however, put pressure on financial intermediation margins, thereby decreasing bank profitability. This could lead to a contraction in bank lending, thereby reducing deposits and hence contracting money supply. Central banks might find themselves in a position where they would like to address the maturity transformation concern without constricting money supply. In this case, if banks were restricted to offering only *zero-interest demand-deposit* accounts, then in response to the central bank's policy actions of flattening the yield curve, it would be profitable at the margin for banks to switch from financing long-maturity credit to financing short-maturity credit. Bank earnings as a result of the policy would be compensated for by price increases on any long-maturity assets sold to the central bank, as well as by earning relatively higher interest on short-maturity credit issued. This would ease any pressure on banks to decrease lending and contract money supply. Apart from maintaining money supply, the central bank's actions would reduce the average maturity of bank assets, making banks less susceptible to unwinding risk at a time when this was of concern to the central bank. Despite the appeal of this strategy, it once again does not require too much imagination to envisage how banks might circumvent the restriction of zero interest rates through adventurous off-balance-sheet activities. It is therefore far from certain that the central bank's actions would have the desired outcome without effective micro-prudential controls.

These examples are not meant as policy prescriptions. Their intention is to illustrate the importance of considering both the transaction services of banks in addition to their financial intermediation services when making policy and regulatory decisions. And of course, neither of these illustrations address the problem of bad credit decisions.

3.13. Conclusions

Empirical evidence indicates that different bank deposit account types have different responses to macroeconomic shocks. Monetary policy innovations elicit a much larger interest rate response and explain a larger proportion of interest rate variance for saving-type accounts than transactional-type accounts, both for corporates and households. Monetary policy innovations also elicit a greater response from corporate transactional-type account balances than corporate short-maturity saving-type account balances (the difference for households is not statistically significant). These results highlight difference in policy transmission across these different account types. In addition, shocks to saving-type account balances have the characteristics of savings shocks—unlike those of transactional-type account balances. These differences suggest that different deposit account types reflect different underlying intermediation services, and that these services have different policy transmission affects.

An intriguing result is that bank balance sheets respond to positive expenditure innovations primarily through *household* transactional-type accounts, but they respond to contractionary monetary policy innovations primarily through *corporate* transactional-type accounts. A possible interpretation for this is that they capture two different documented macroeconomic phenomena in South Africa over the period. First, positive expenditure shocks along with rising household deposit balances may capture rising government expenditure on municipal wages during this period. Second, positive monetary policy shocks along with decreasing corporate deposit balances may reflect the substantial fall in private investment expenditure over this period. More research is required to unpack this result.

Results support an underlying premise of this study: that the liabilities of bank balance sheets reflect both financial and monetary intermediation services. Additional work could investigate whether responses found in this study are asymmetric to positive and negative policy shocks, or whether there are other nonlinearities to the effects. Future research could also include data on bank assets, and incorporate more explicit measures of international financial factors. The role of bank liabilities as a medium-of-exchange is a unique and defining characteristic of banks, yet it is largely absent from macroeconomic models. The results of this study suggest that this problematic. Transactional-type account balances comprise nearly 25% of South African bank deposit liabilities, but there is a dearth of literature focused on understanding how these transaction-focused accounts affect policy transmission in South Africa or elsewhere.

3.14. Appendices

3.14.1. Appendix 3A: Results reported in referenced studies

Figure 3-A.1: Results reported in Bernanke & Blinder (1992)

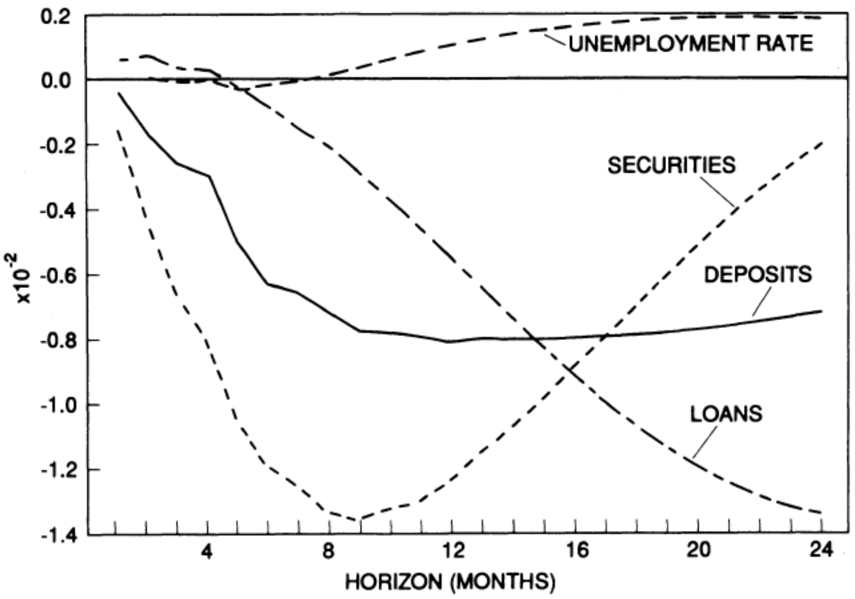
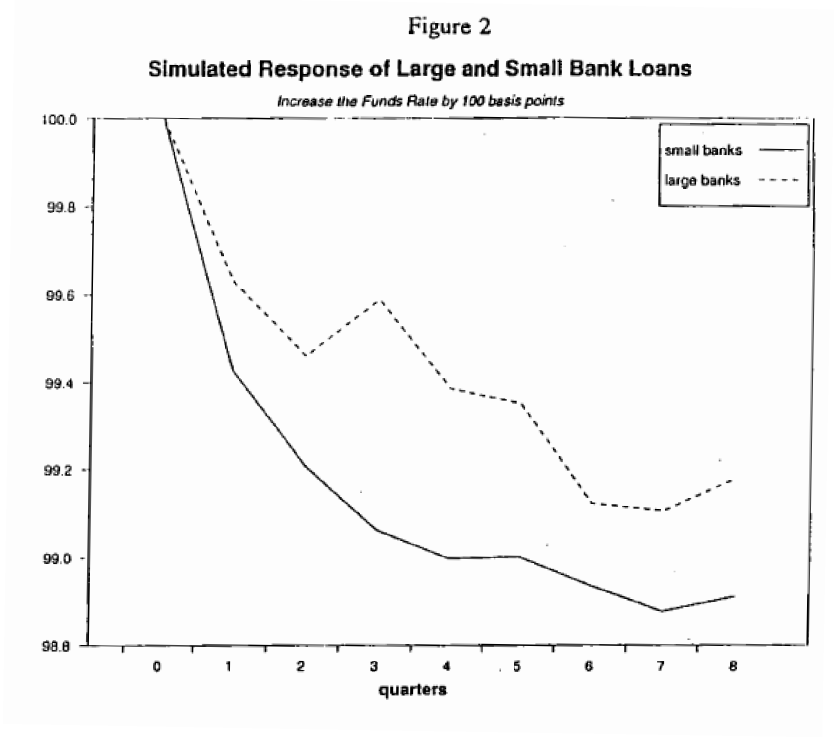


FIGURE 4. RESPONSES TO A SHOCK TO THE FUNDS RATE

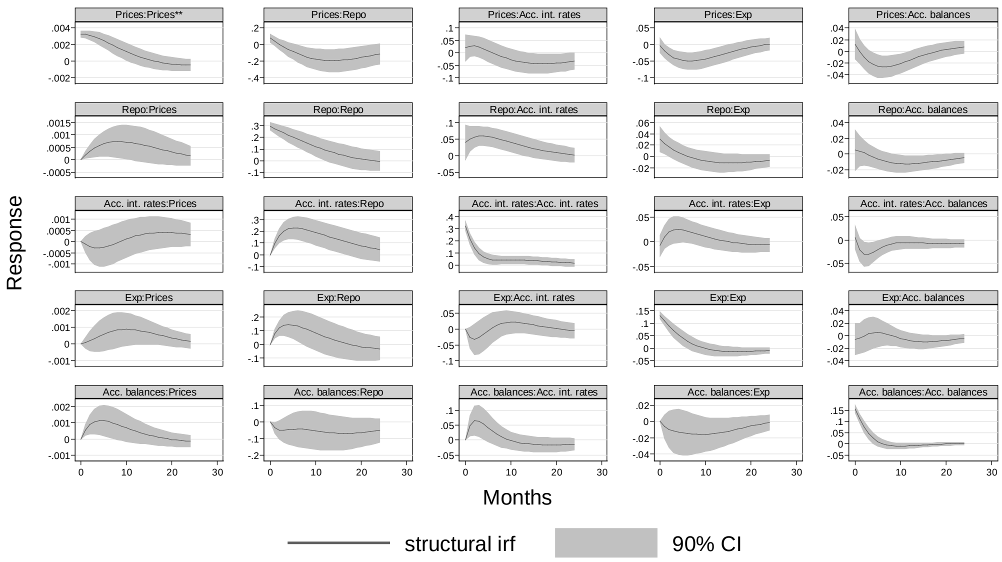
Figure A.2: Results reported in Kashyap & Stein (1995)



3.14.2. *Appendix 3B: Corporate results*

Figure 3-B.1: Corporate IRFs and FEVDs

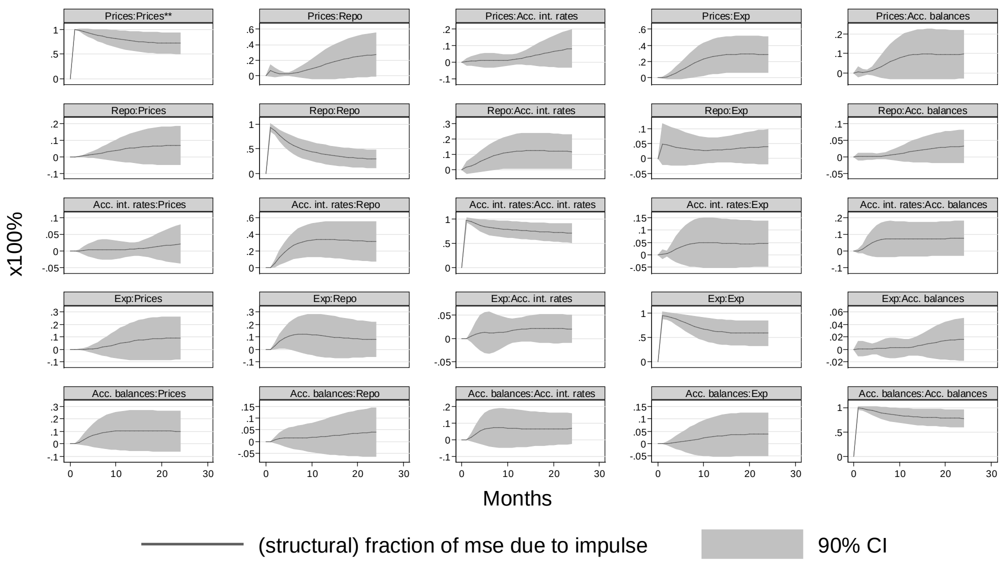
(i) Corporate cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

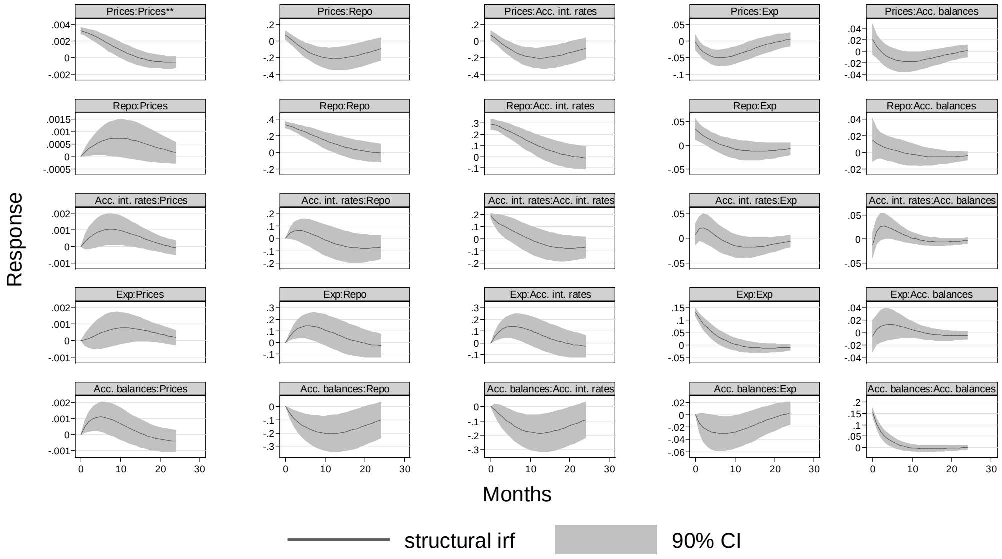
(ii) Forecast-error variance decomposition of corporate cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

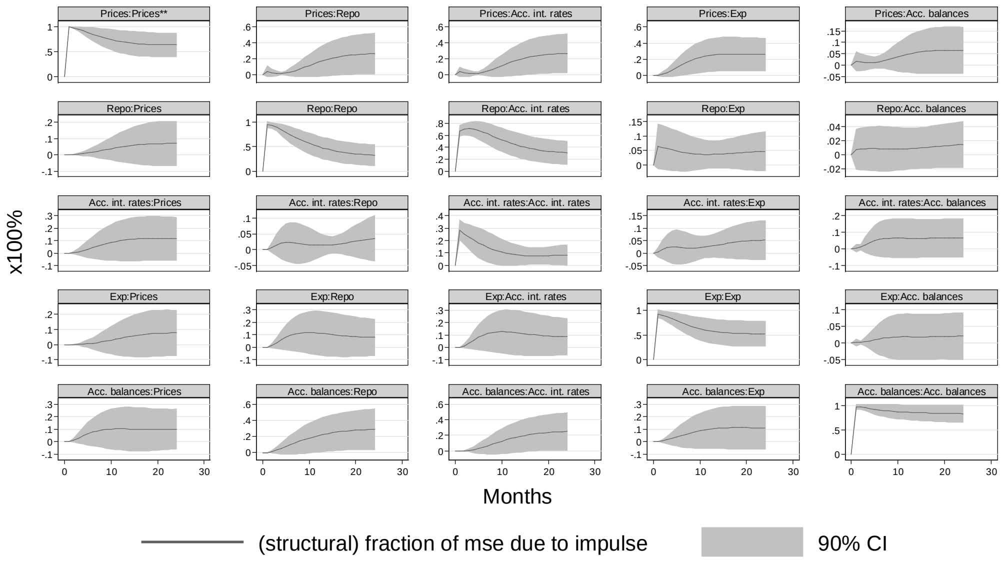
(iii) Corporate call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

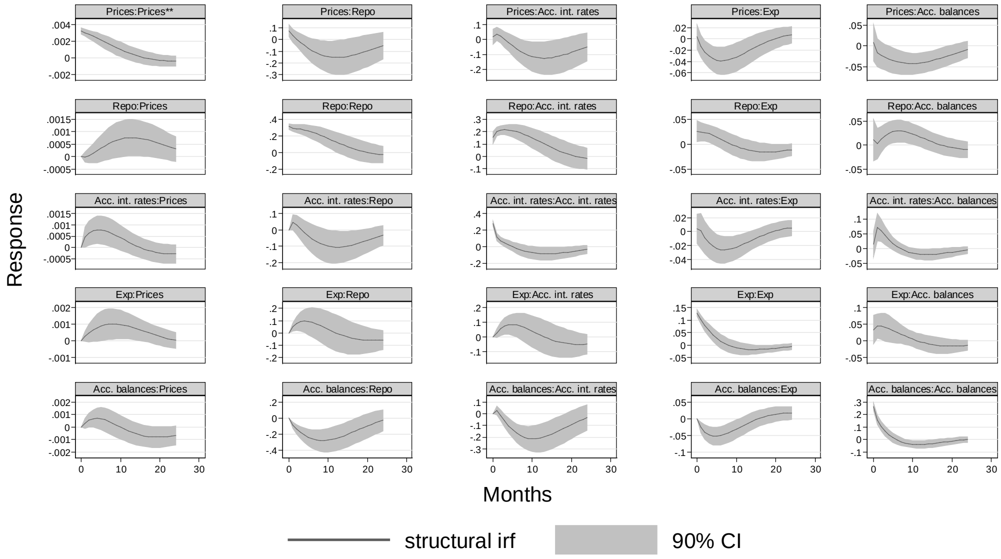
(iv) Forecast-error variance decomposition of corporate call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

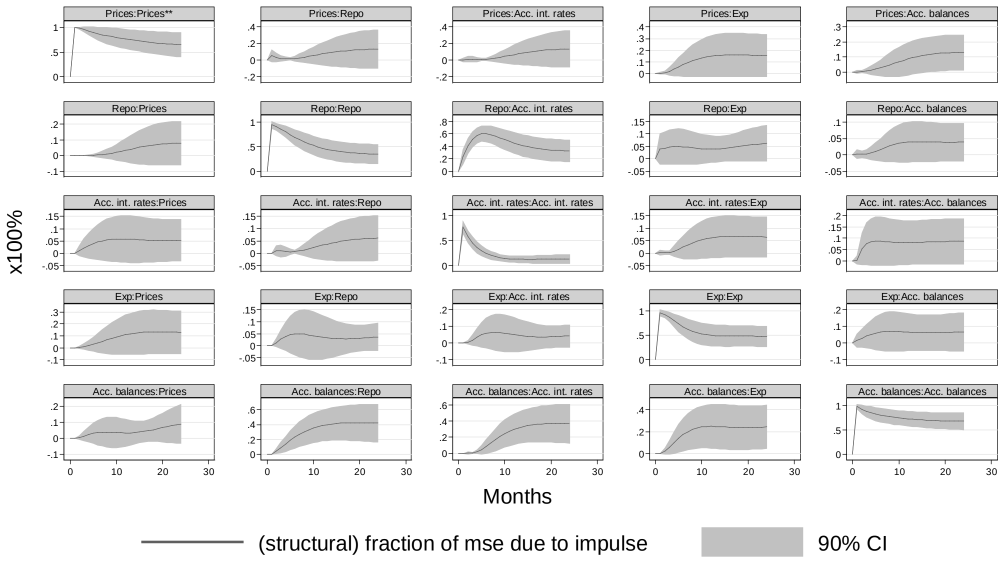
(v) Corporate 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

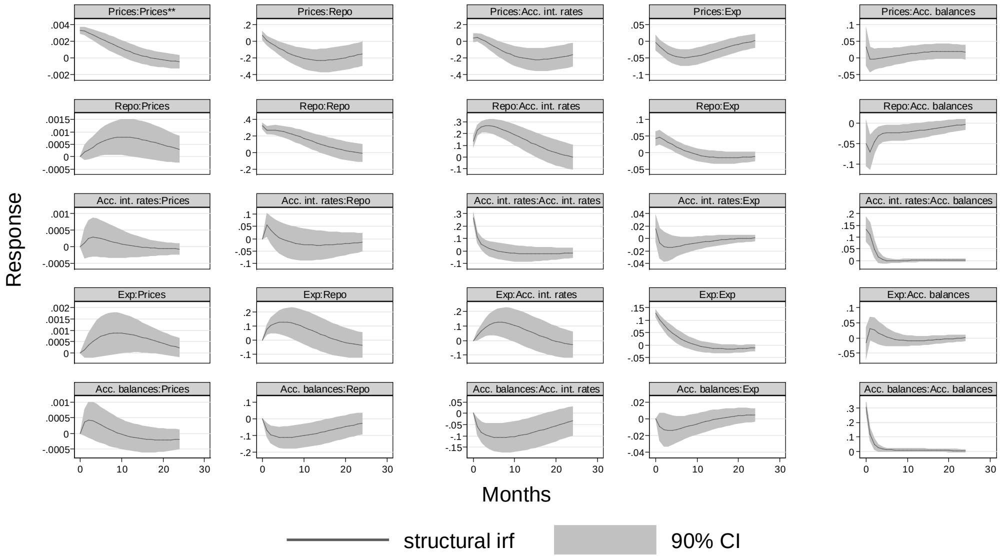
(vi) FEVD of corporate 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

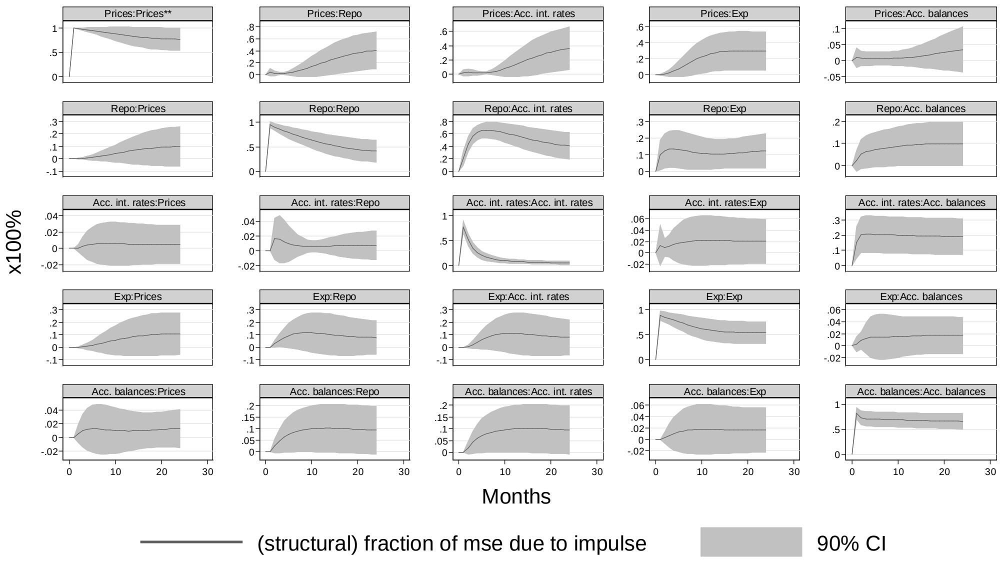
(vii) Corporate 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

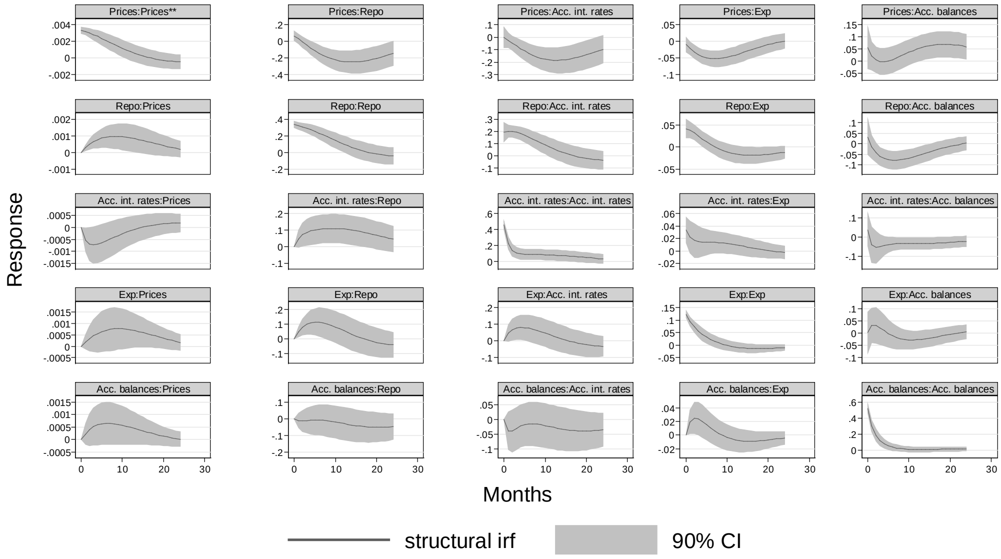
(viii) FEVD of corporate 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

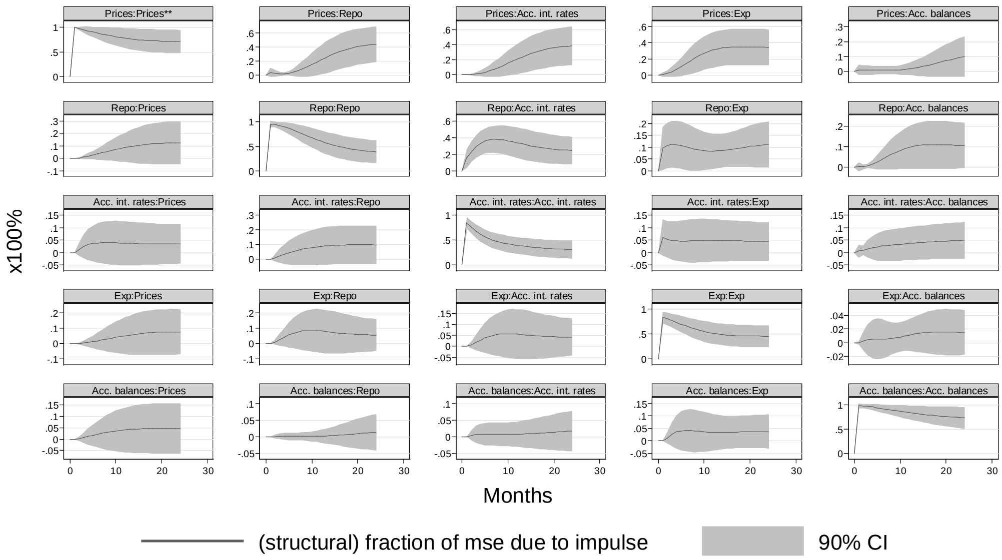
(ix) Corporate 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

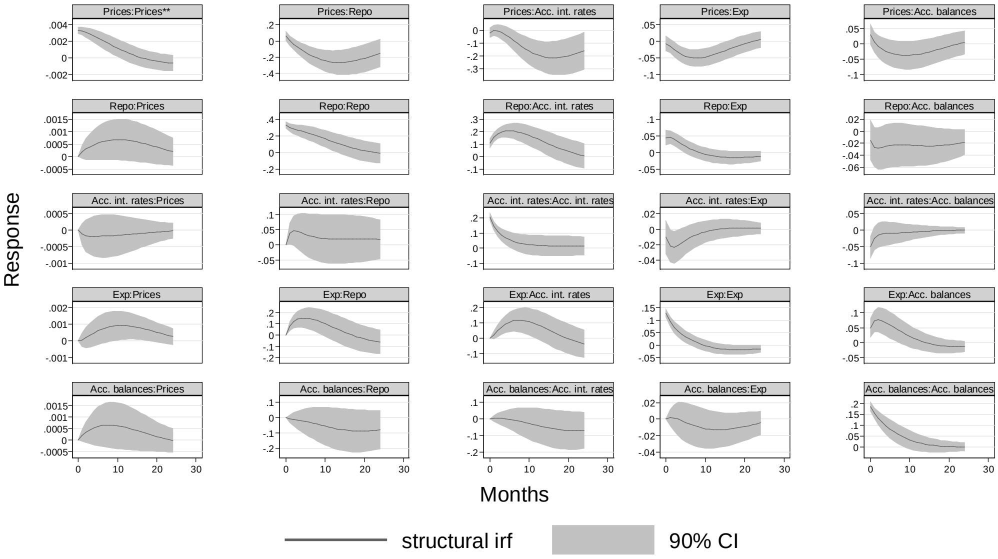
(x) FEVD of corporate 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

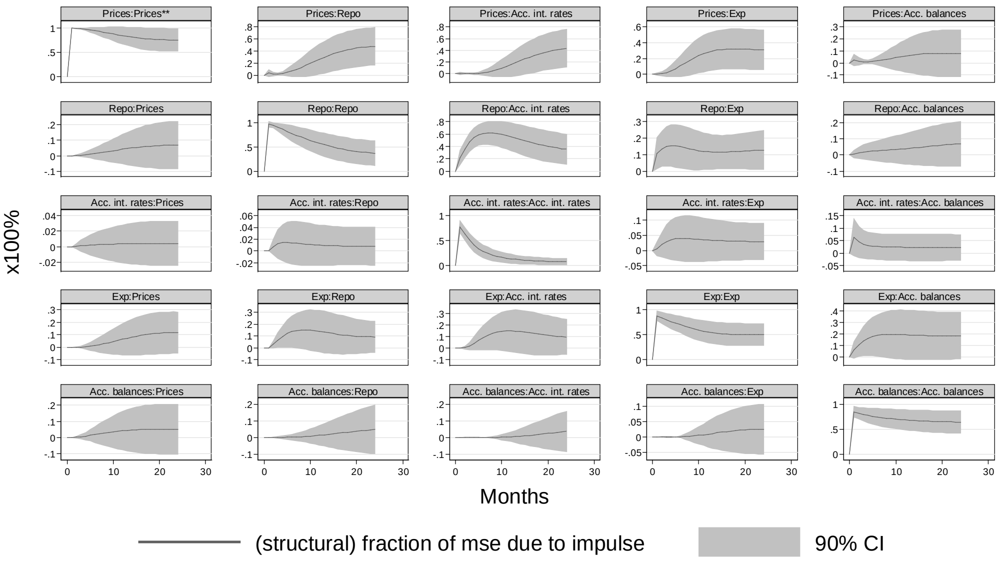
(xi) Corporate 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

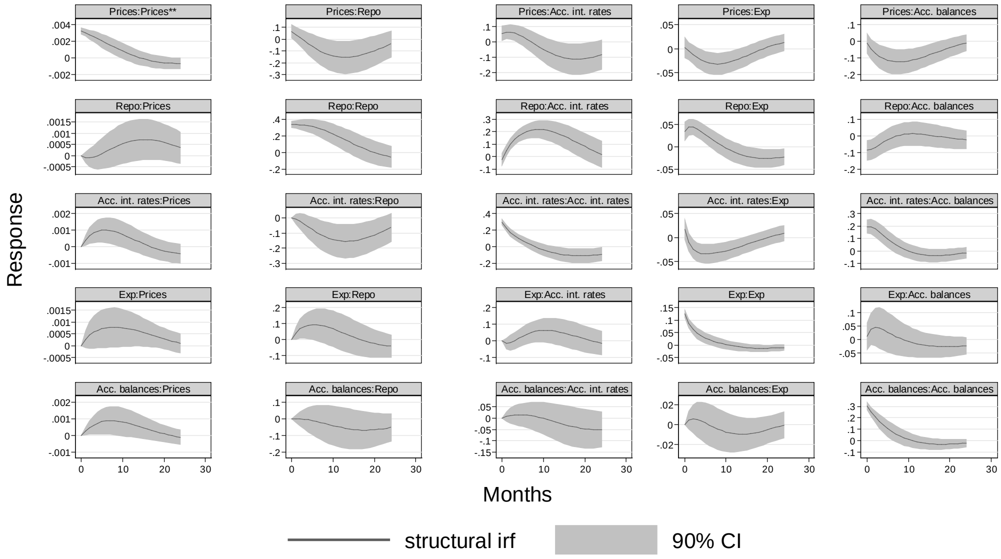
(xii) FEVD of corporate 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

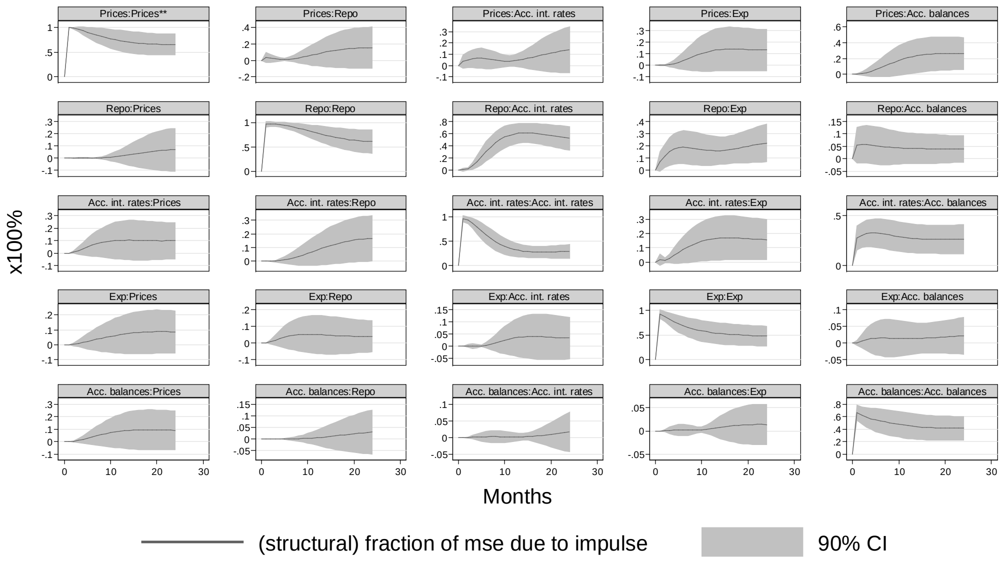
(xiii) Corporate 1-3yr account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

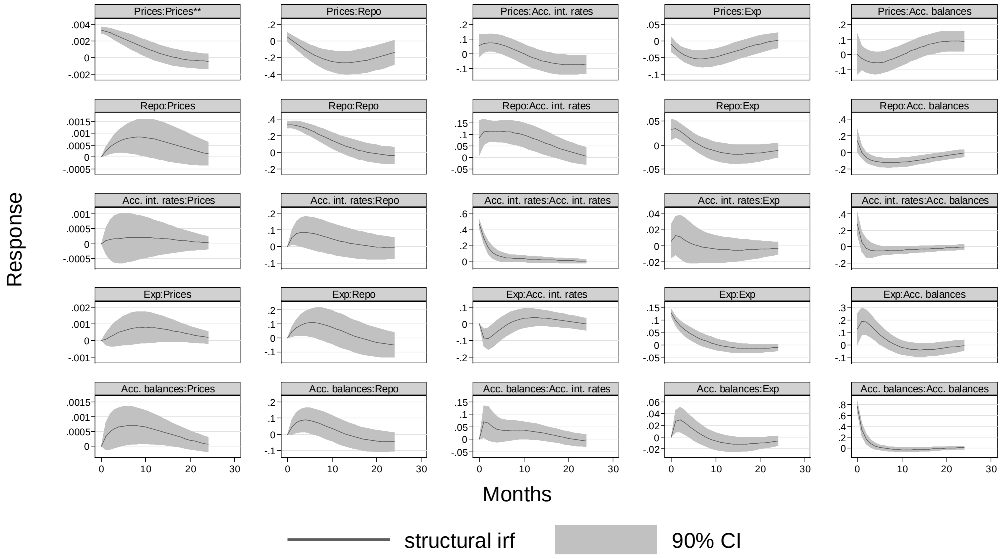
(xiv) FEVD of corporate 1-3y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

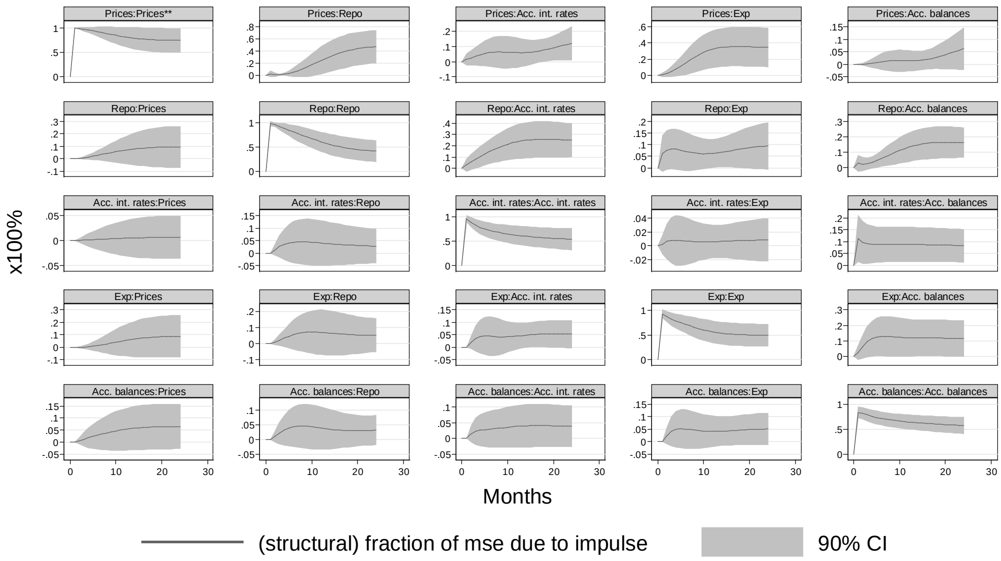
(xv) Corporate 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

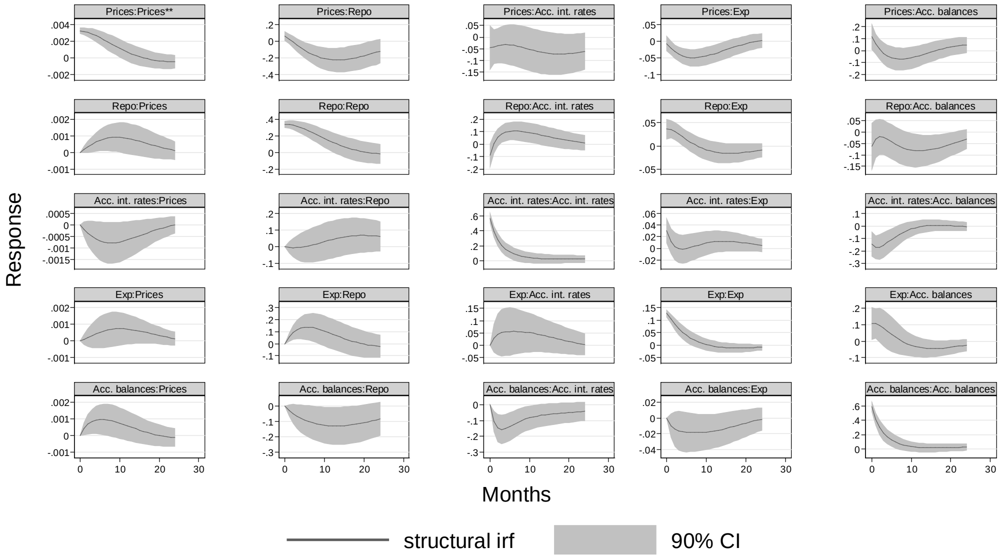
(xvi) FEVD of corporate 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

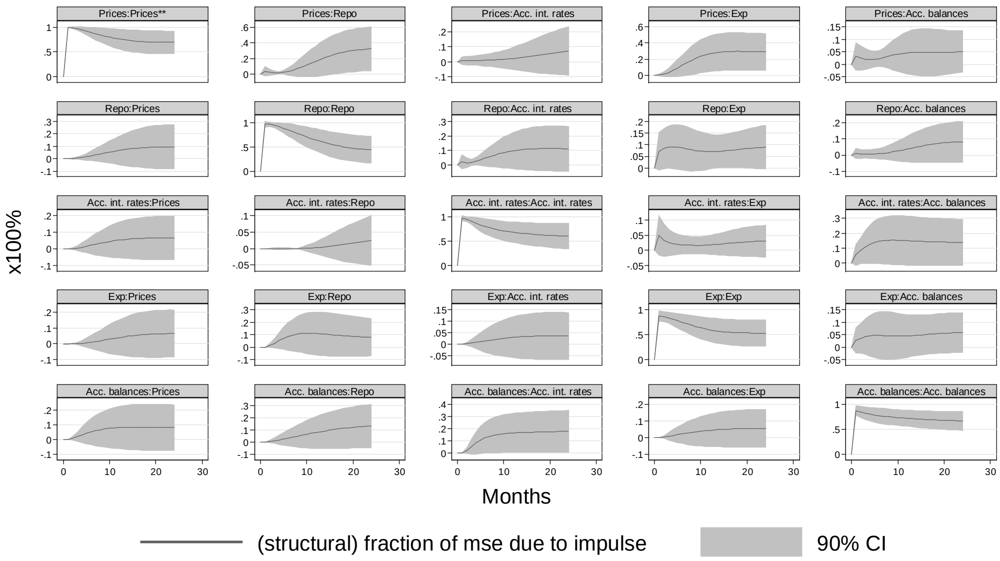
(xvii) Corporate 5y+ account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

(xviii) FEVD of corporate 5y+ account IRFs*



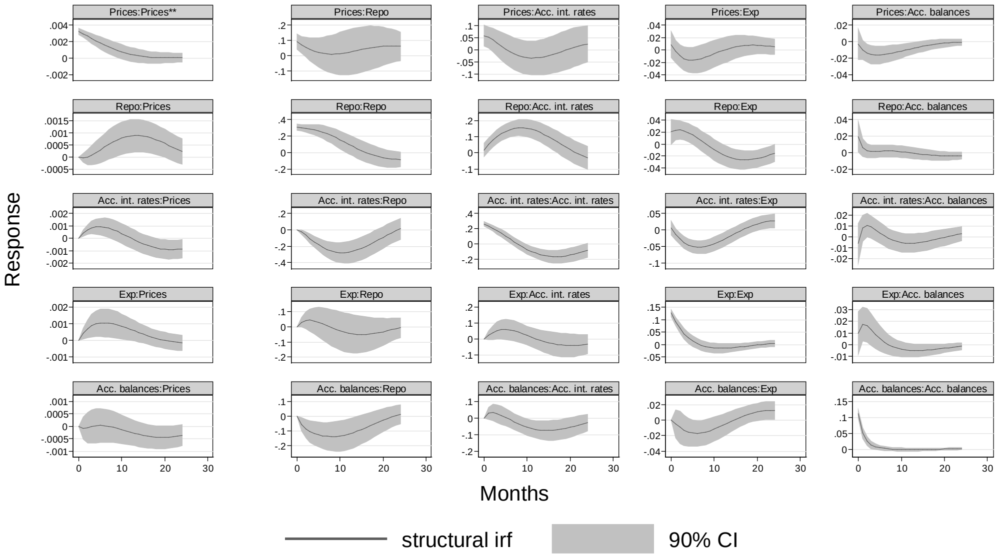
Notes:

* Response to a single std. dev. shock
** Impulse:Response

3.14.3. Appendix 3C: Household results

Figure 3-C.1: Household IRFs and FEVDs

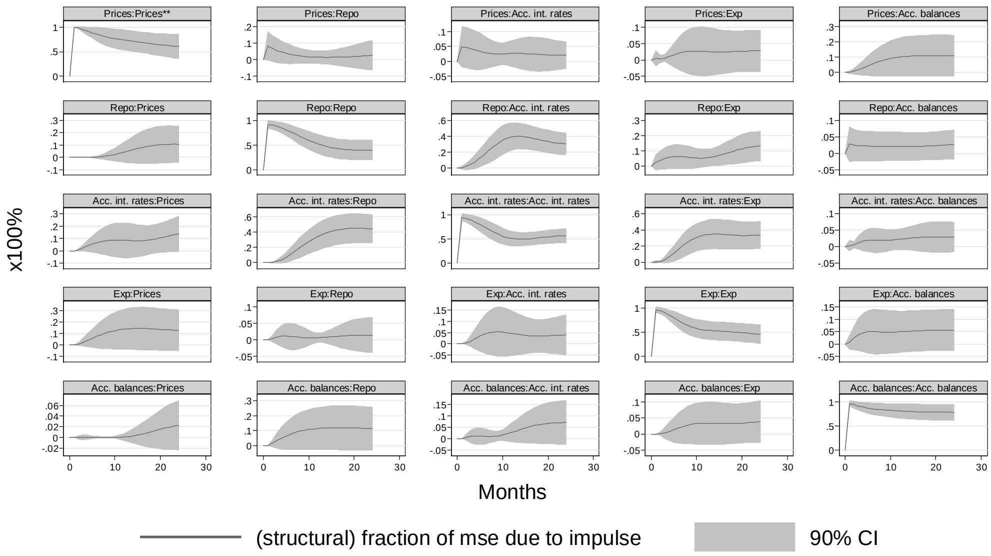
(i) Household cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

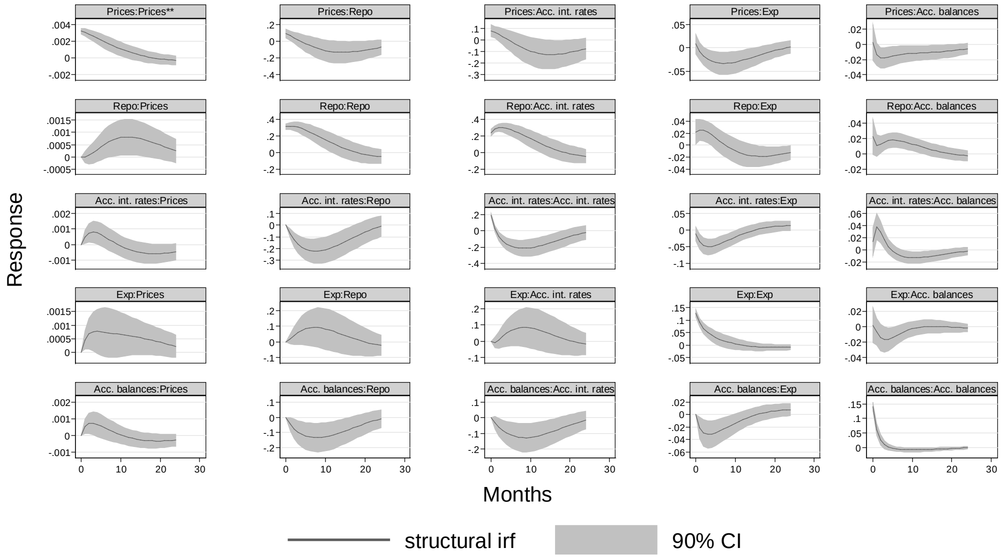
(ii) Forecast-error variance decomposition of household cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

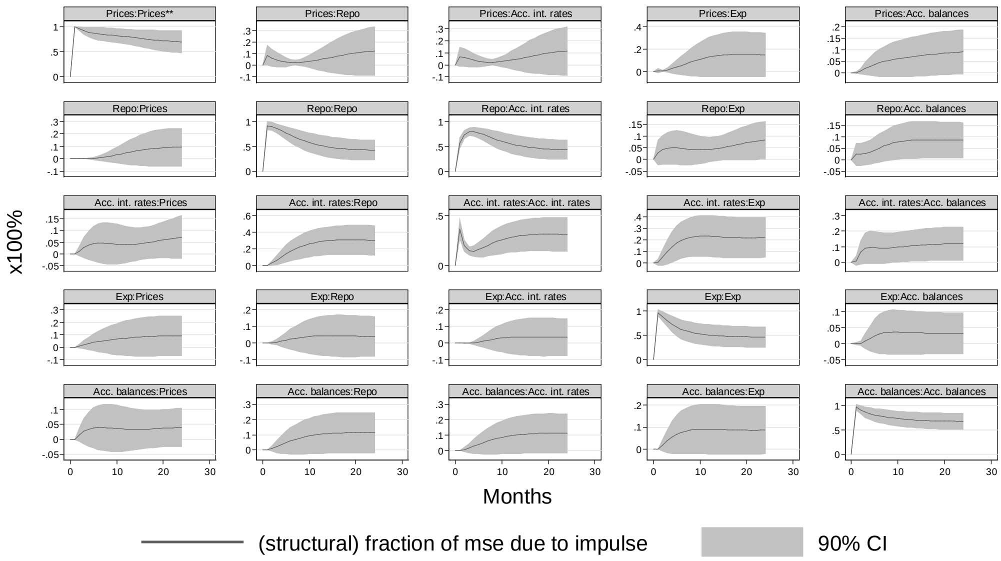
(iii) Household call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

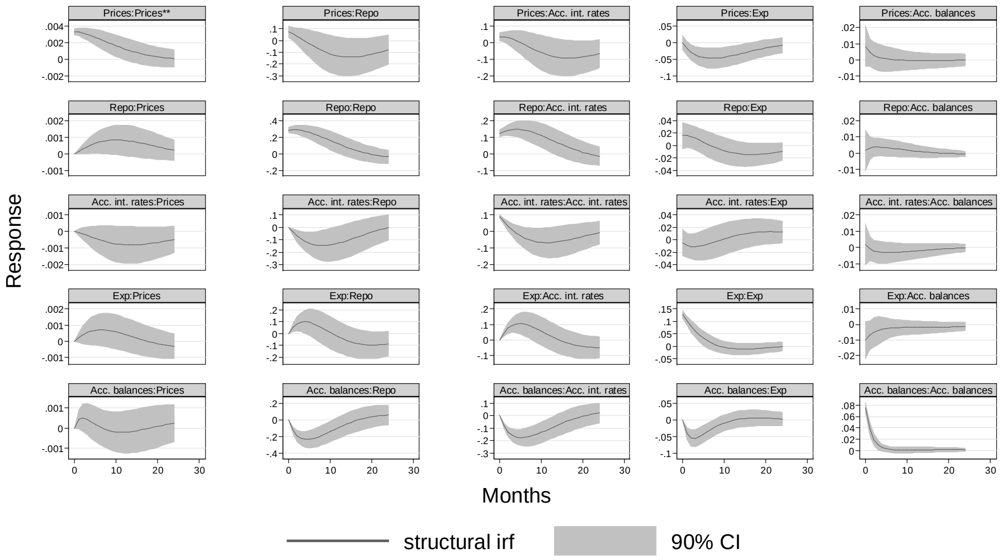
(iv) Forecast-error variance decomposition of household call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

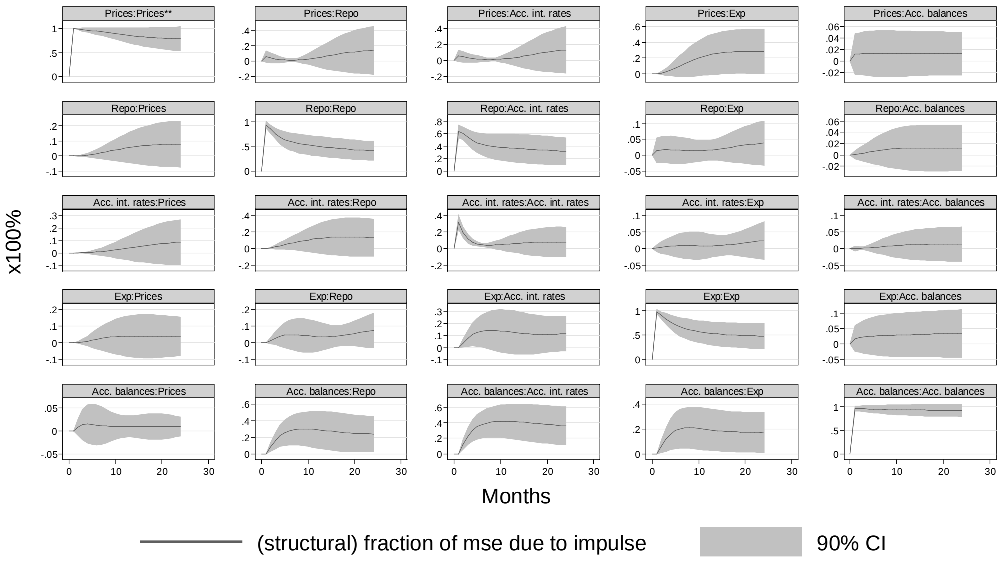
(v) Household savings account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

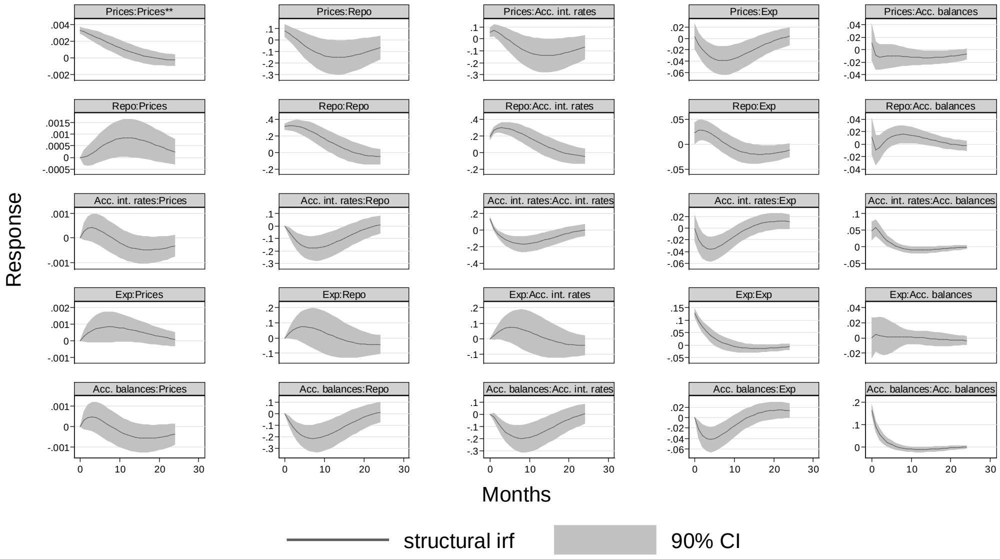
(vi) Forecast-error variance decomposition of household savings account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

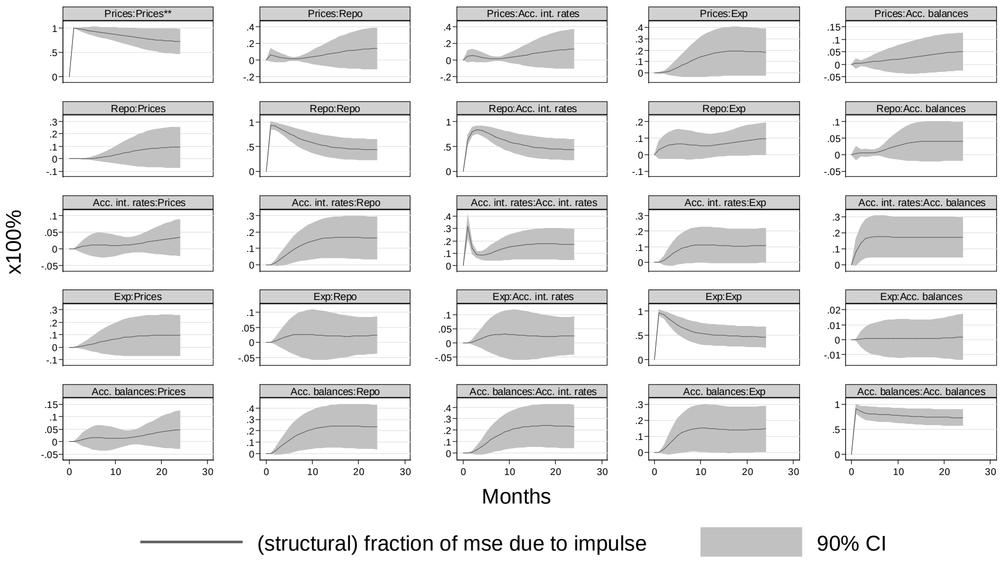
(vii) Household 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

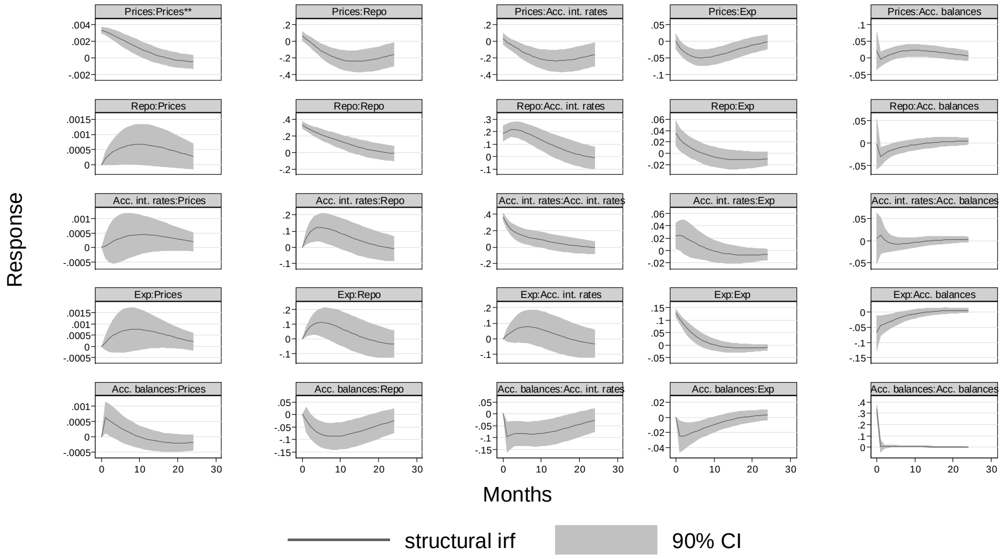
(viii) FEVD of household 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

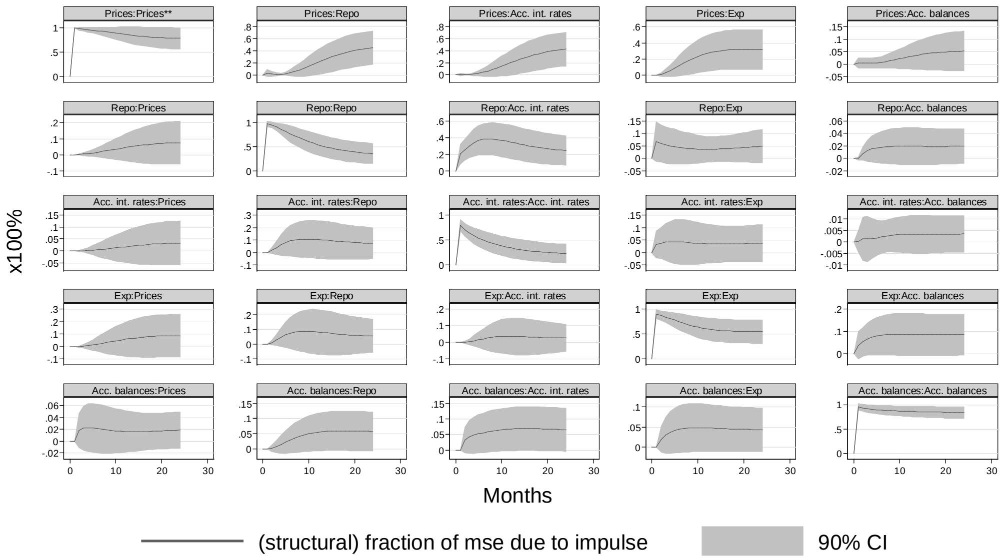
(ix) Household 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

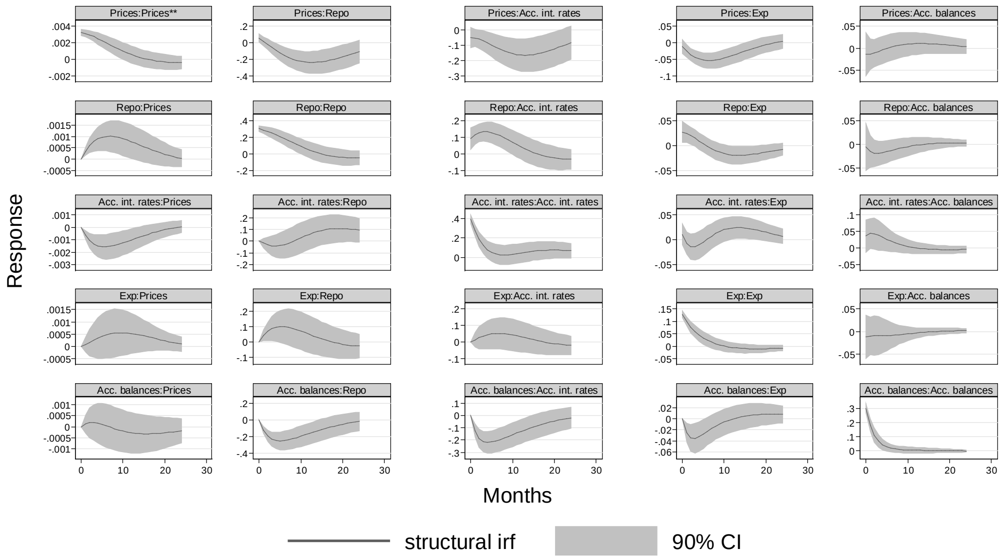
(x) FEVD of household 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

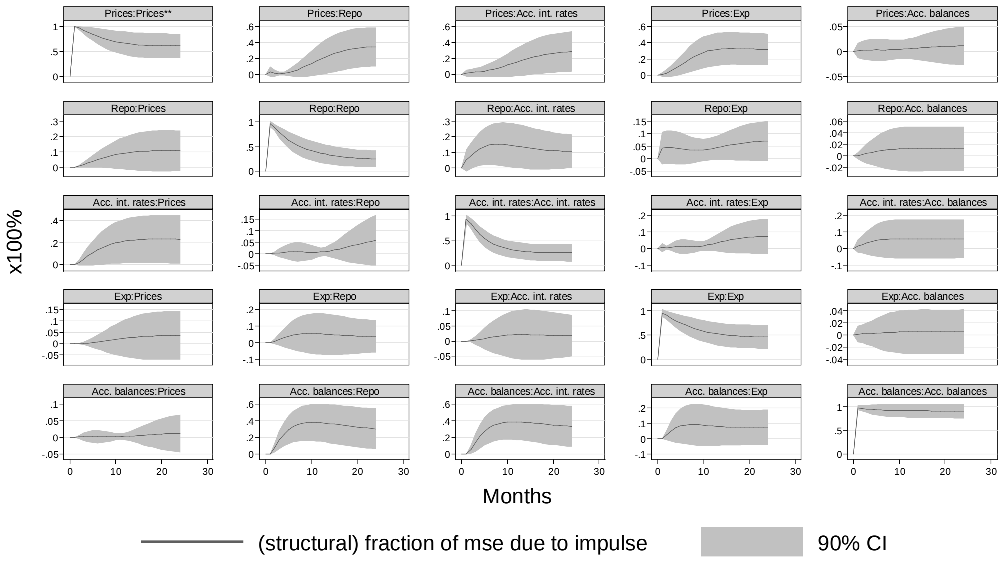
(xi) Household 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

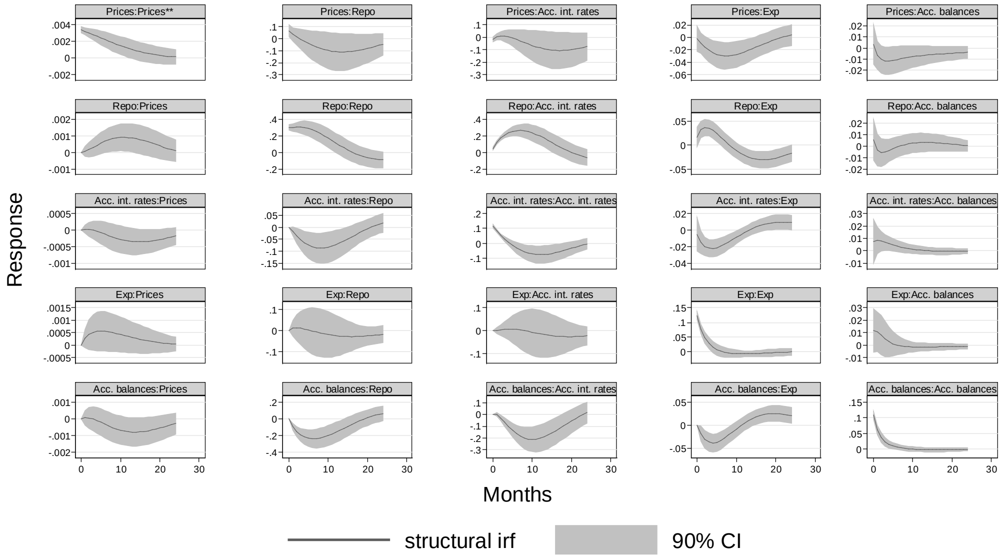
(xii) FEVD of household 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

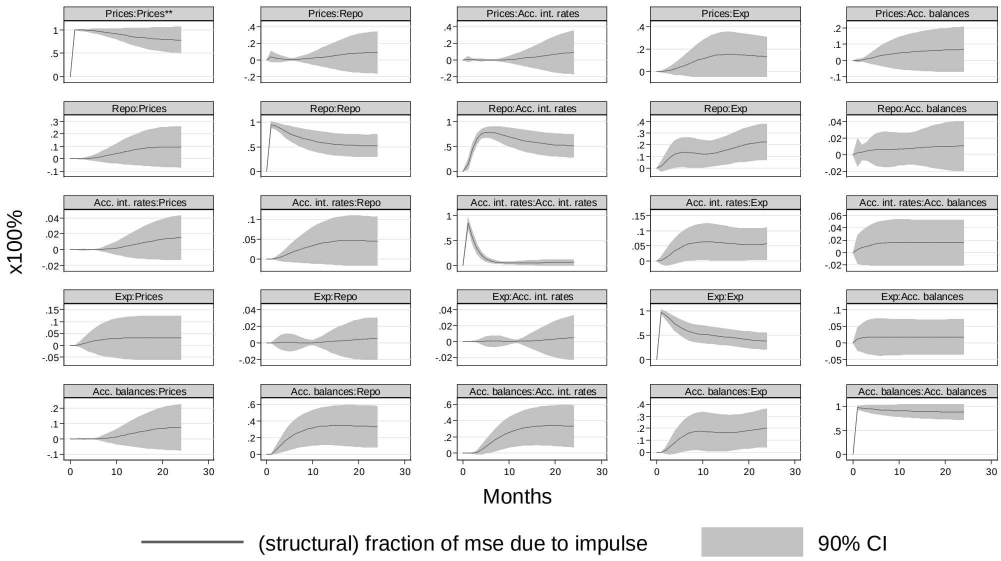
(xiii) Household 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

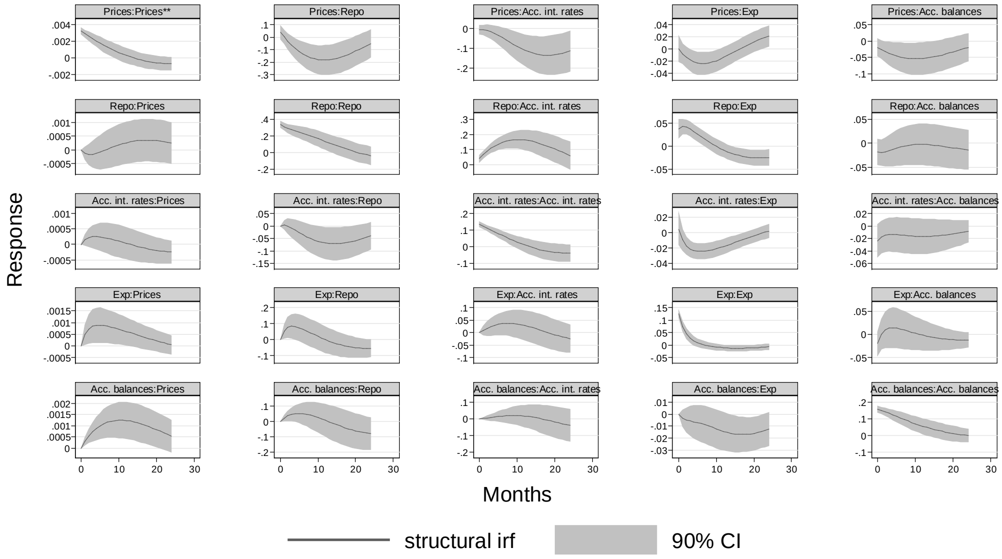
(xiv) FEVD of household 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

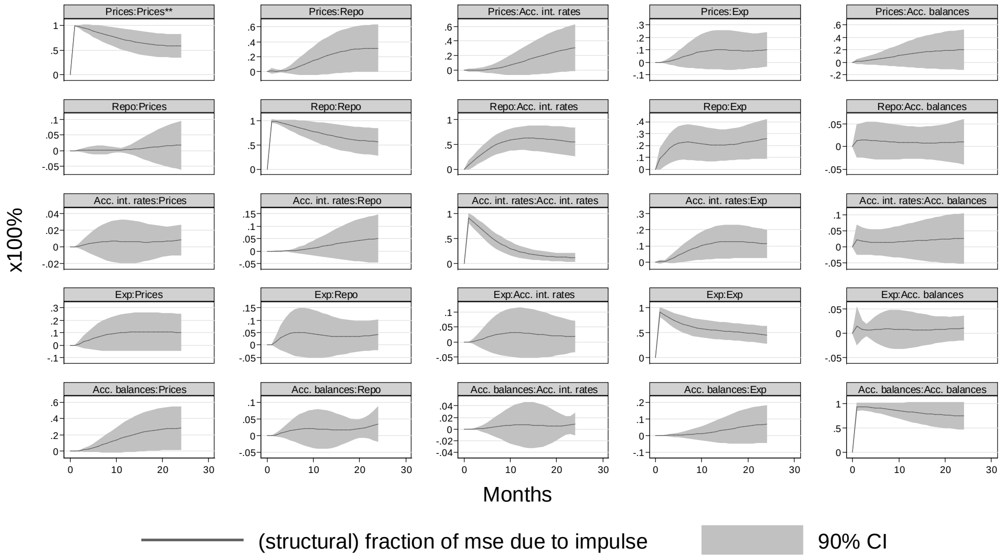
(xv) Household 1-3y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

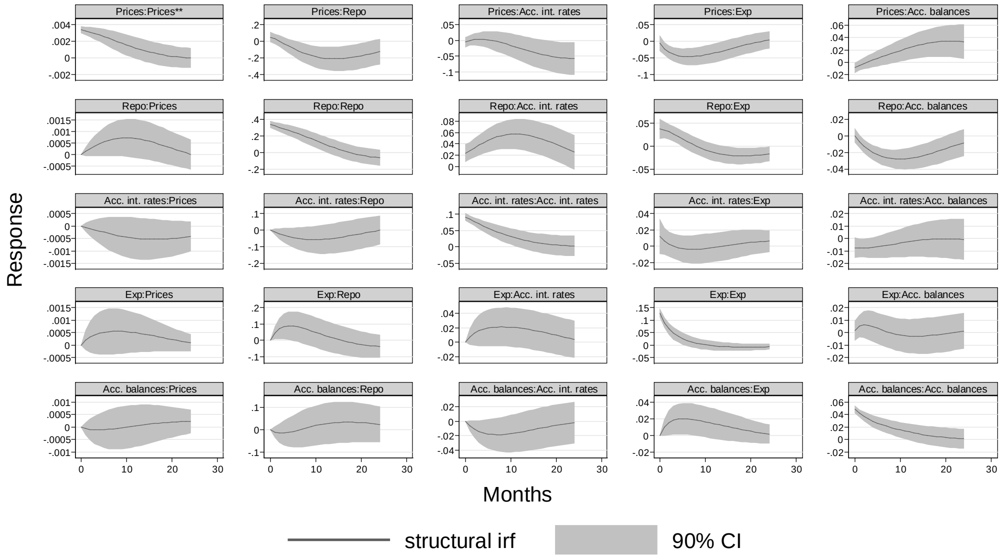
(xvi) FEVD of household 1-3y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

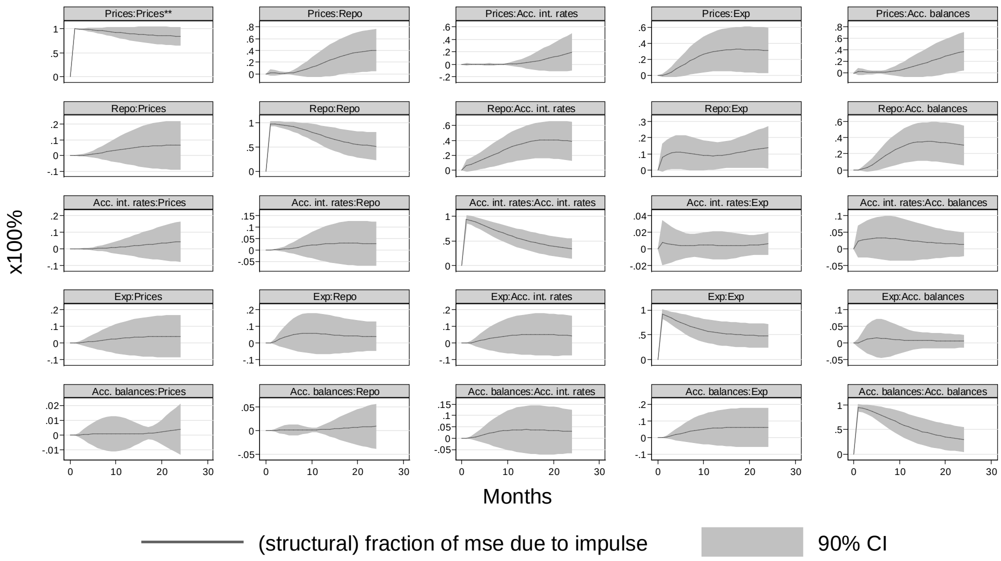
(xvii) Household 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

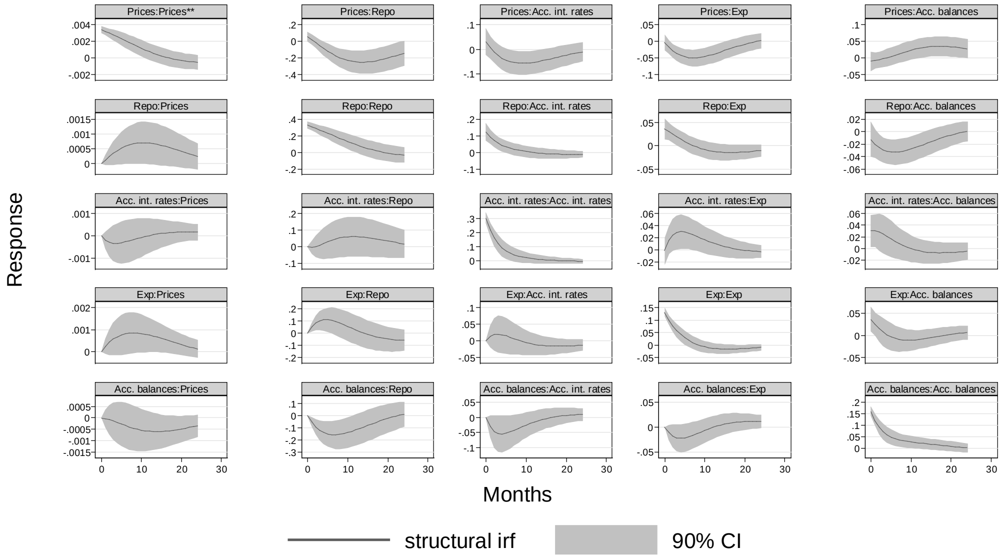
(xviii) FEVD of household 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

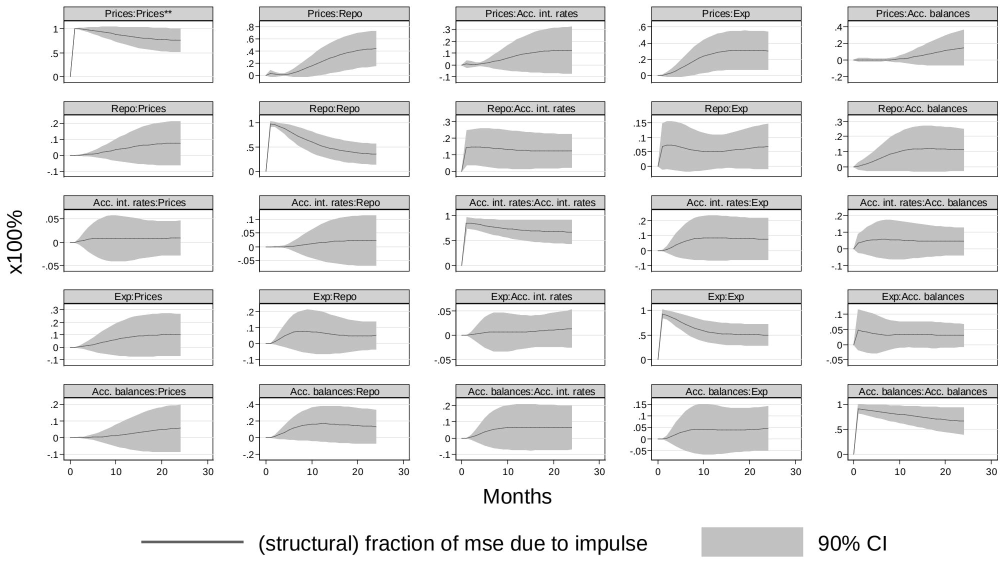
(xix) Household 5y+ account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

(xx) FEVD of household 5y+ account IRFs*



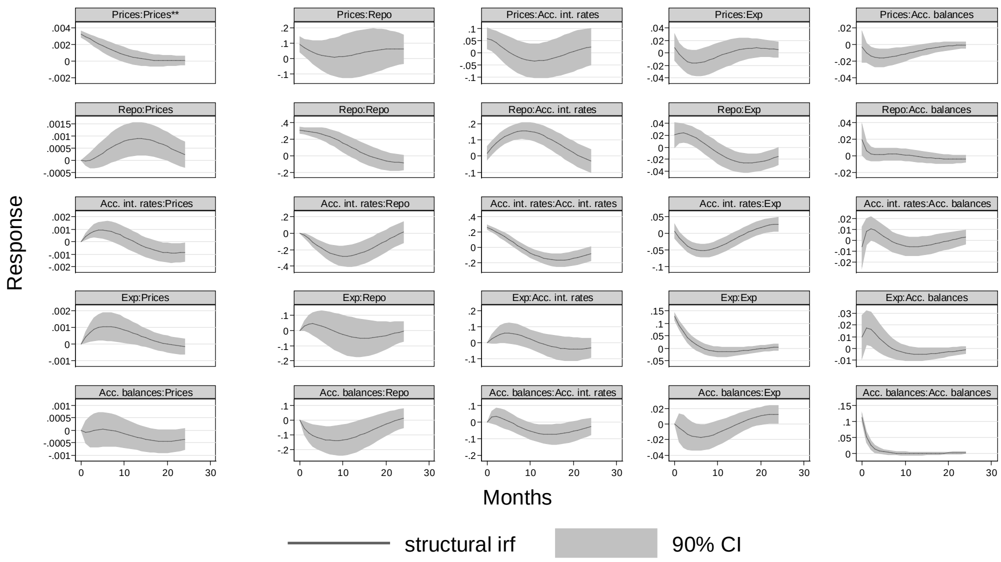
Notes:

* Response to a single std. dev. shock
** Impulse:Response

3.14.4. Appendix 3D: Selected robustness test results

Figure 3-D.1: Corporate IRFs and FEVDs

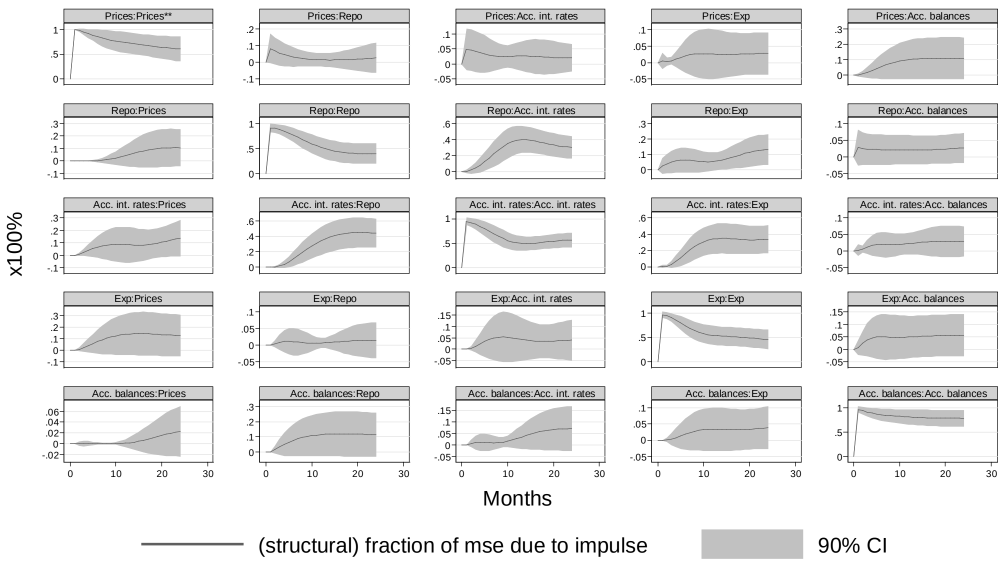
(i) Household cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

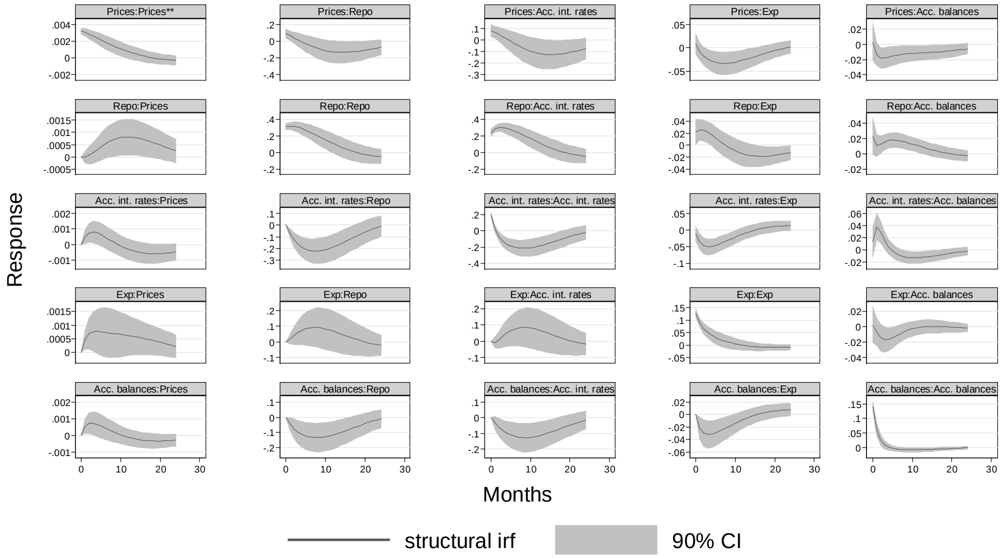
(ii) Forecast-error variance decomposition of household cheque account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

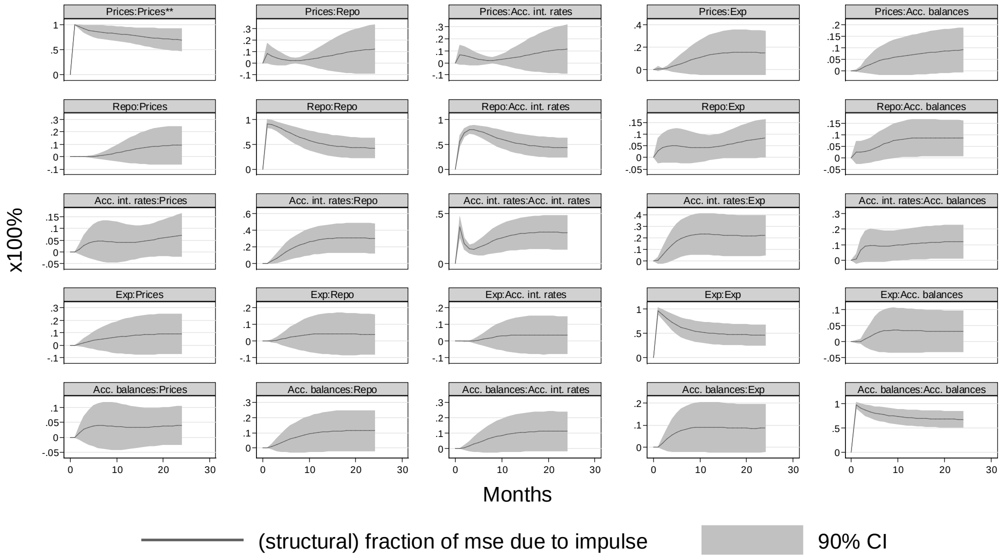
(iii) Household call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

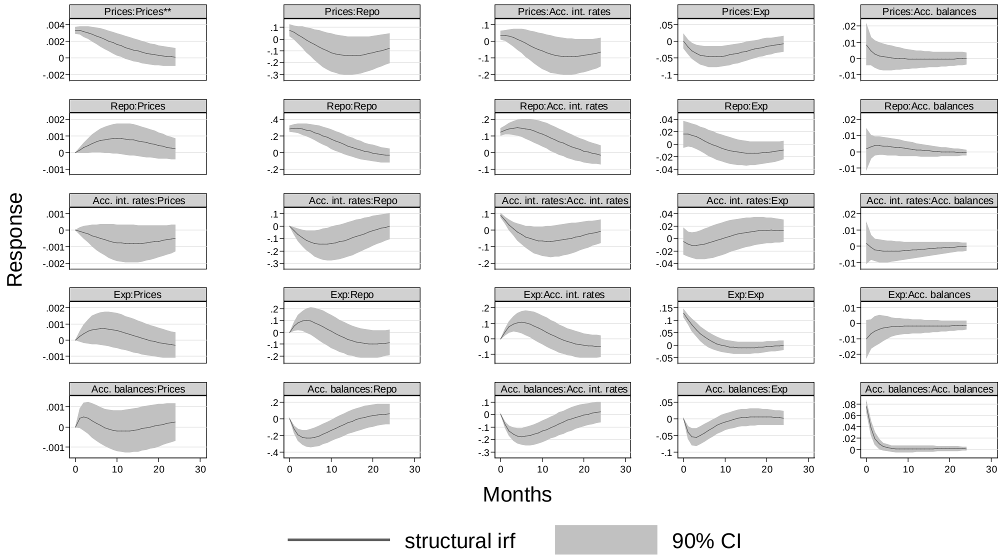
(iv) Forecast-error variance decomposition of household call account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

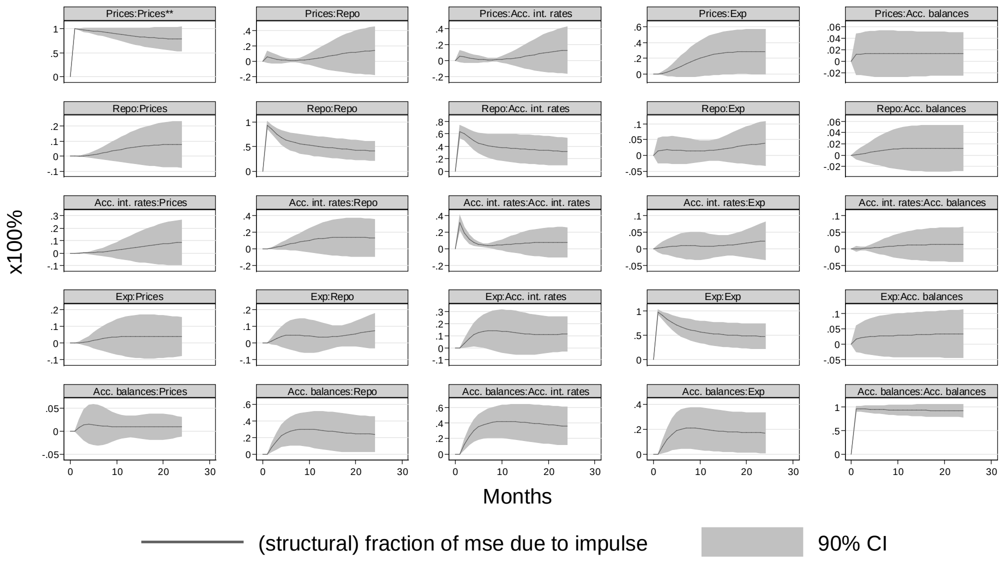
(v) Household savings account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

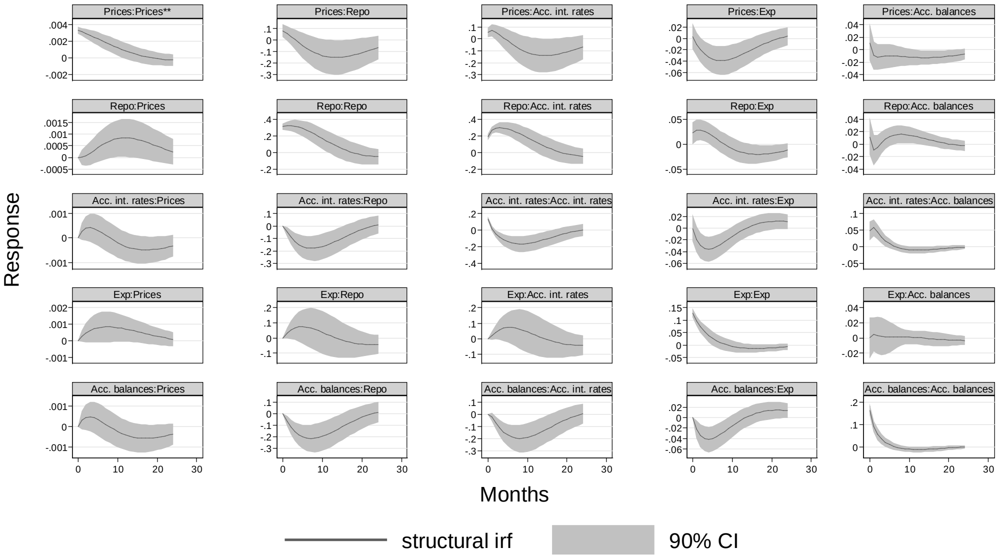
(vi) Forecast-error variance decomposition of household savings account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

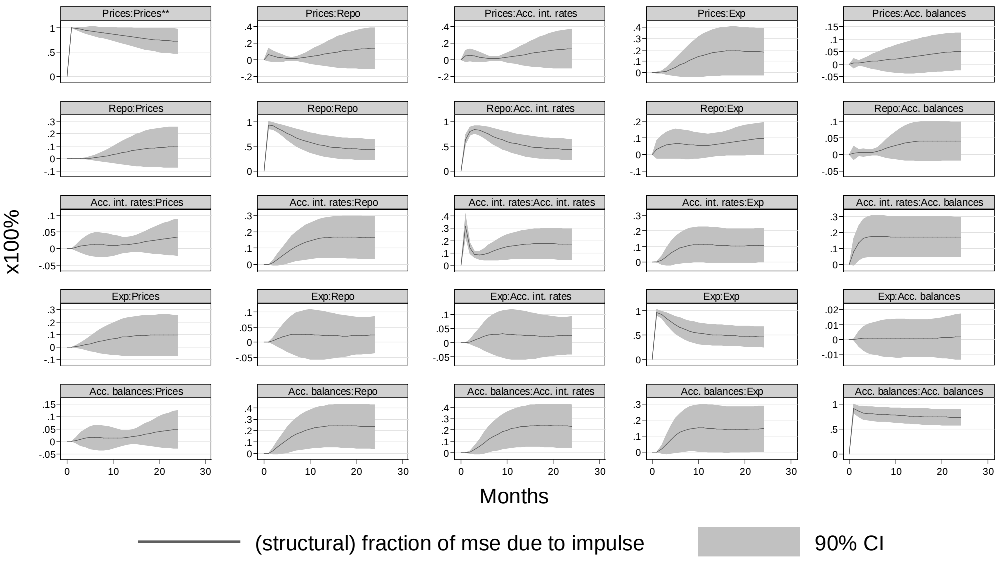
(vii) Household 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

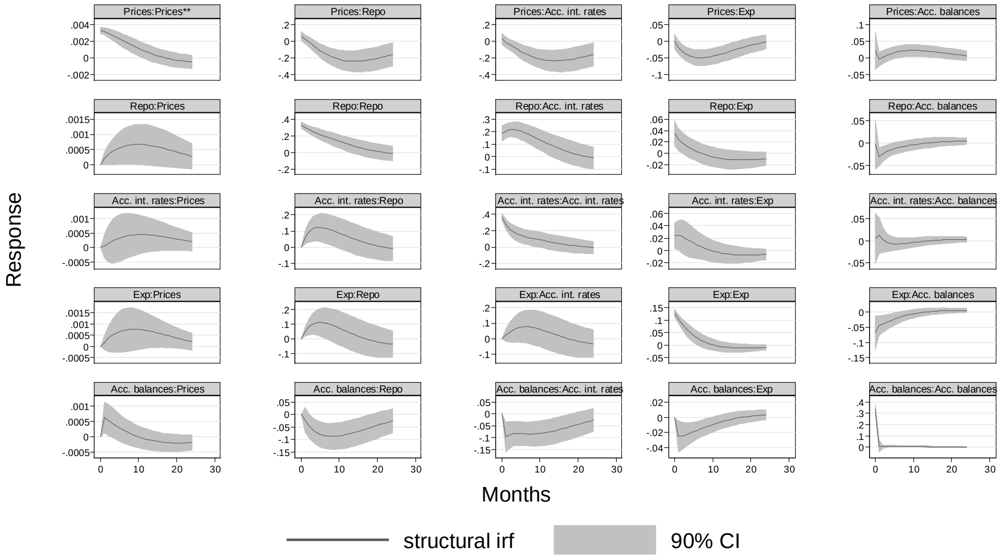
(viii) FEVD of household 32d account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

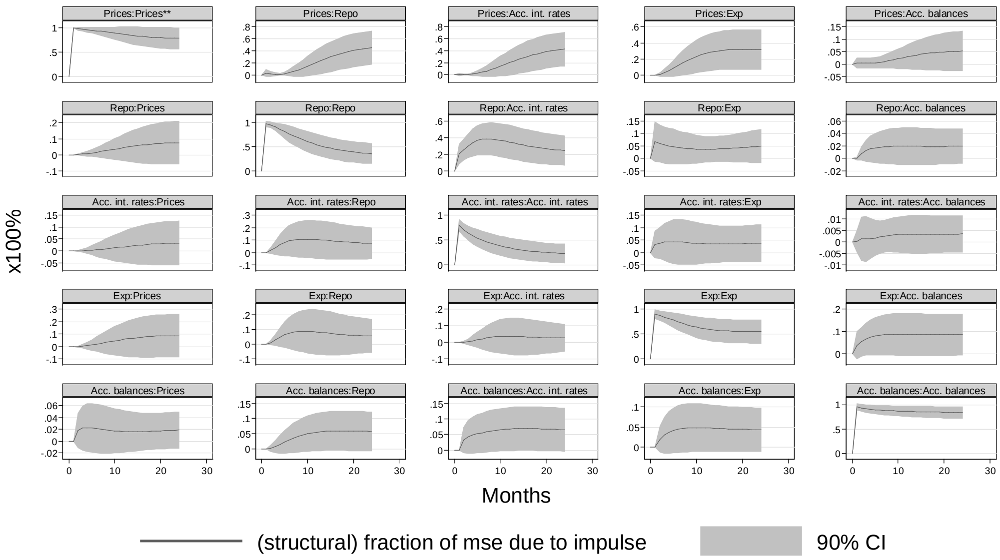
(ix) Household 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

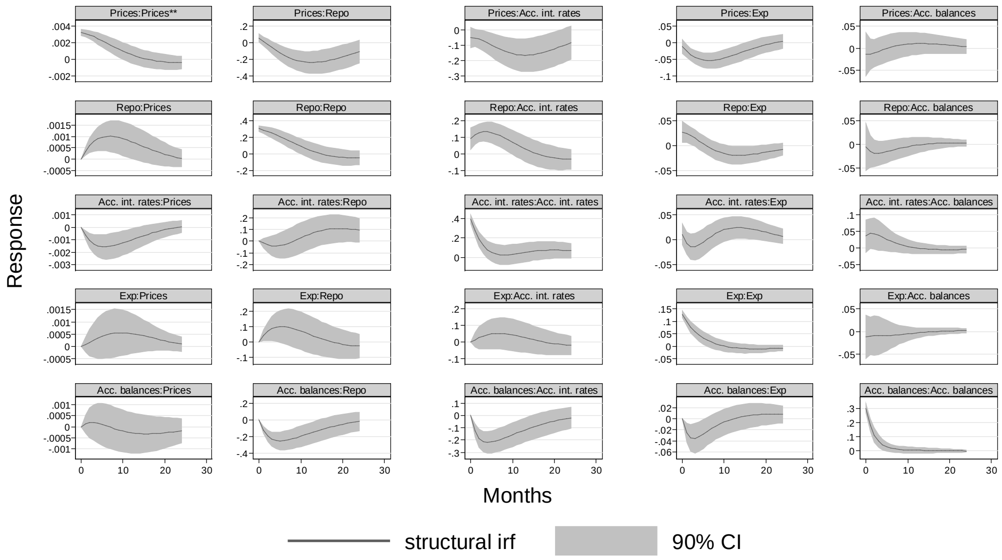
(x) FEVD of household 3m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

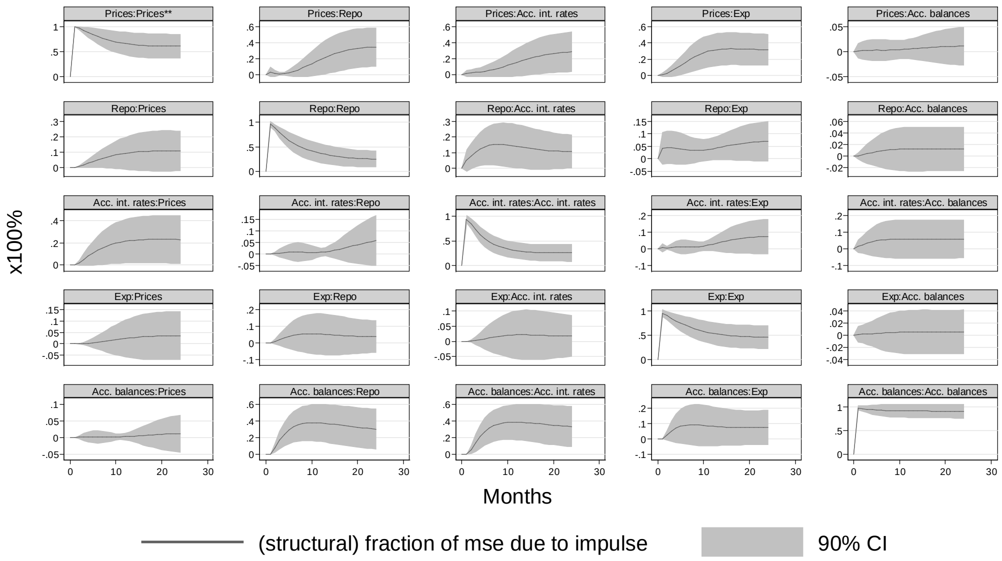
(xi) Household 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

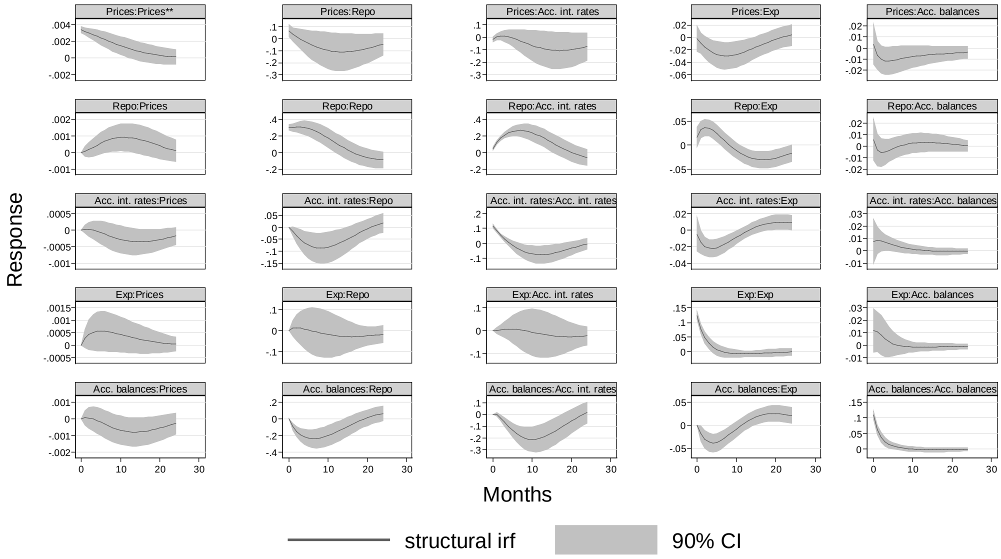
(xii) FEVD of household 6m account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

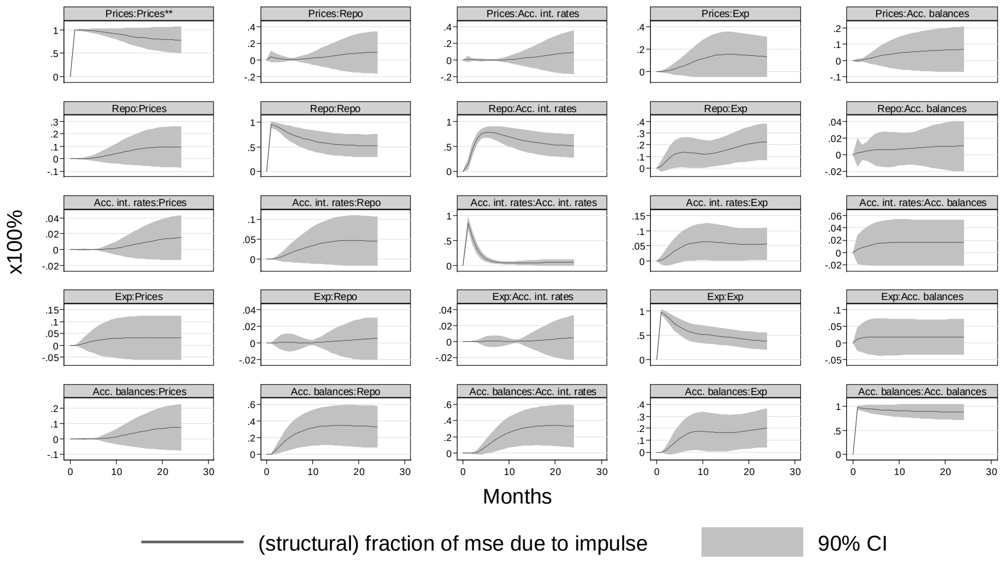
(xiii) Household 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

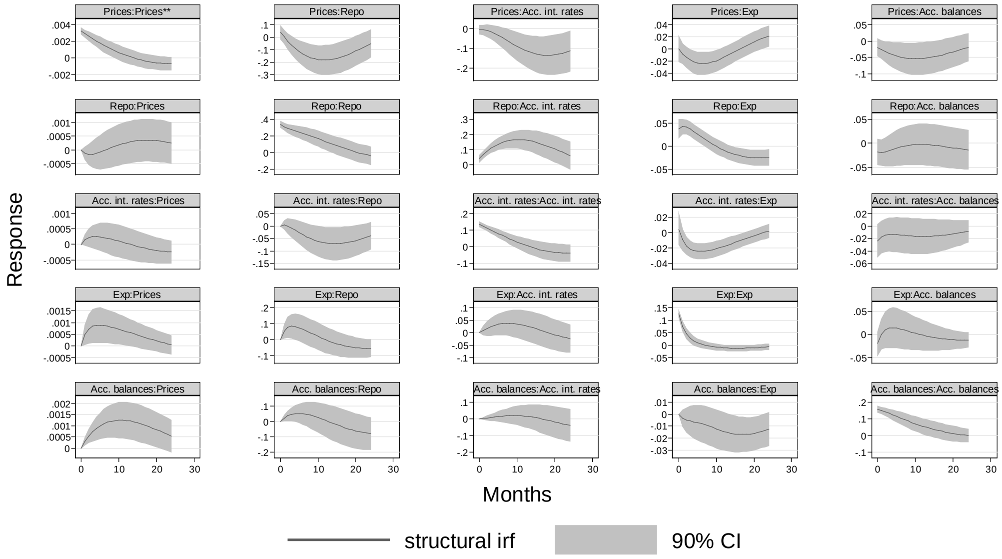
(xiv) FEVD of household 1y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

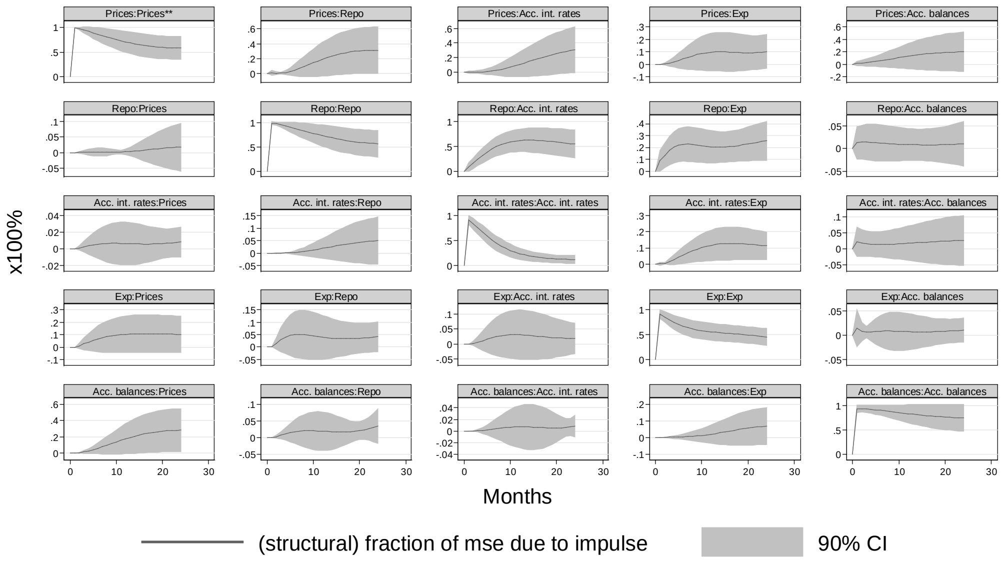
(xv) Household 1-3y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

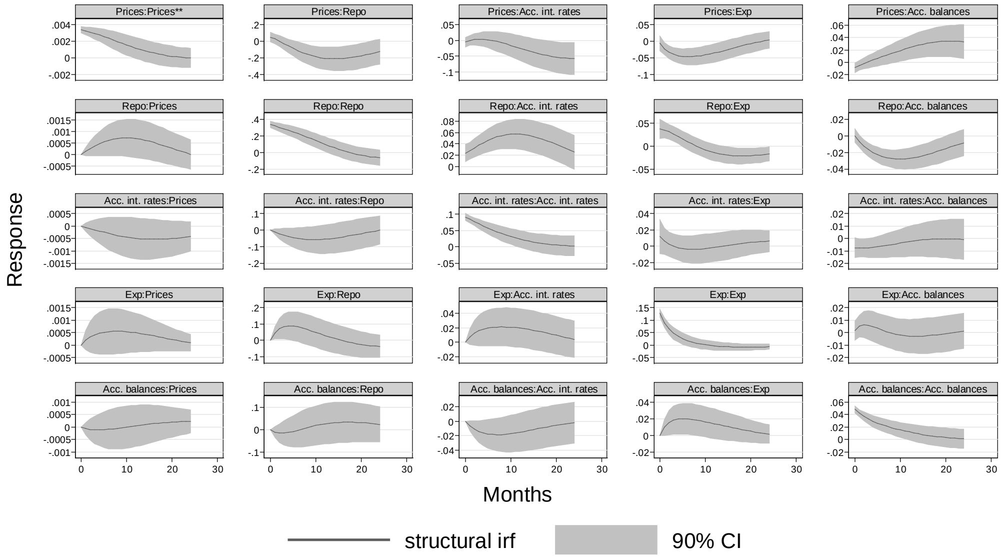
(xvi) FEVD of household 1-3y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

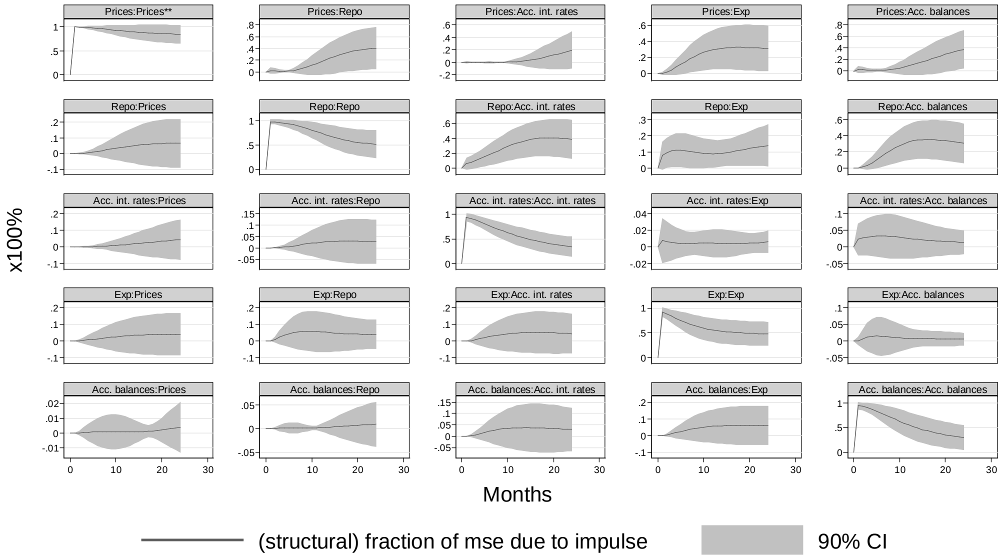
(xvii) Household 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

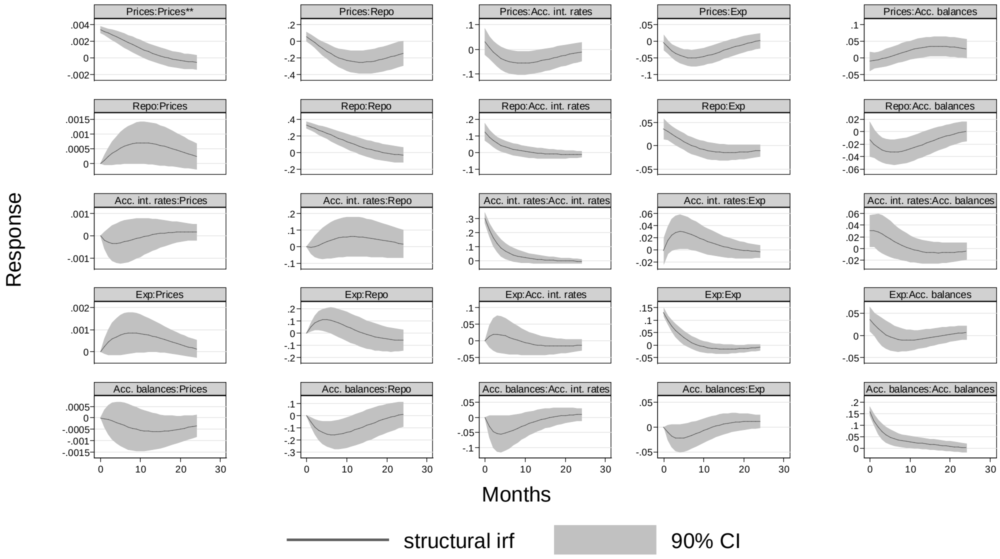
(xviii) FEVD of household 3-5y account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

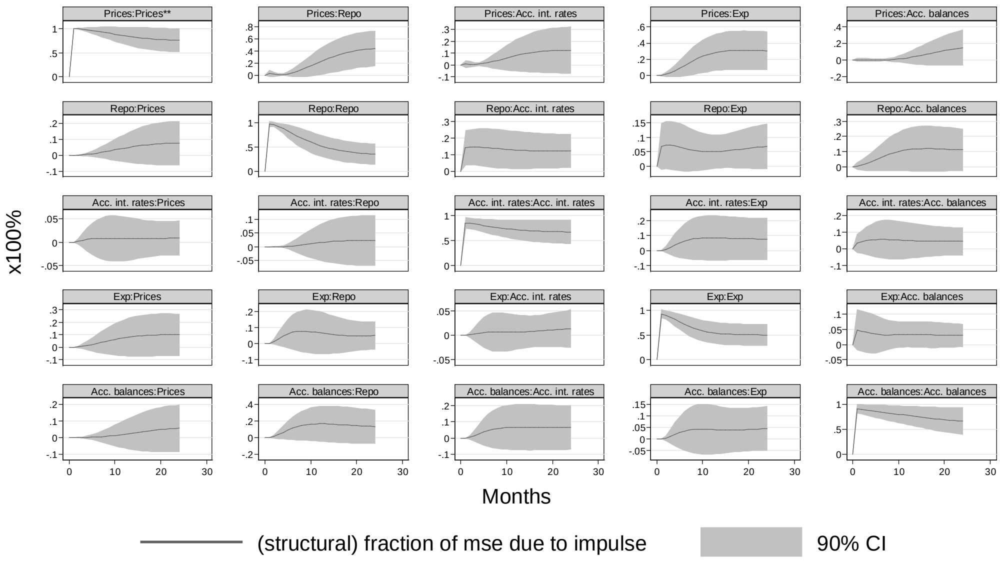
(xix) Household 5y+ account IRFs*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

(xx) FEVD of household 5y+ account IRFs*

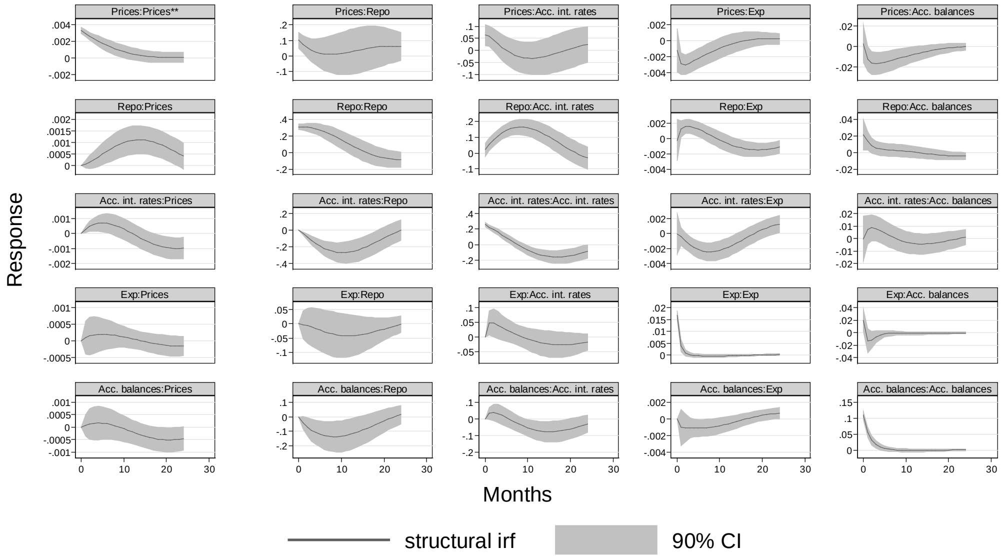


Notes:

* Response to a single std. dev. shock
** Impulse:Response

Figure D.2: Household IRFs and FEVDs

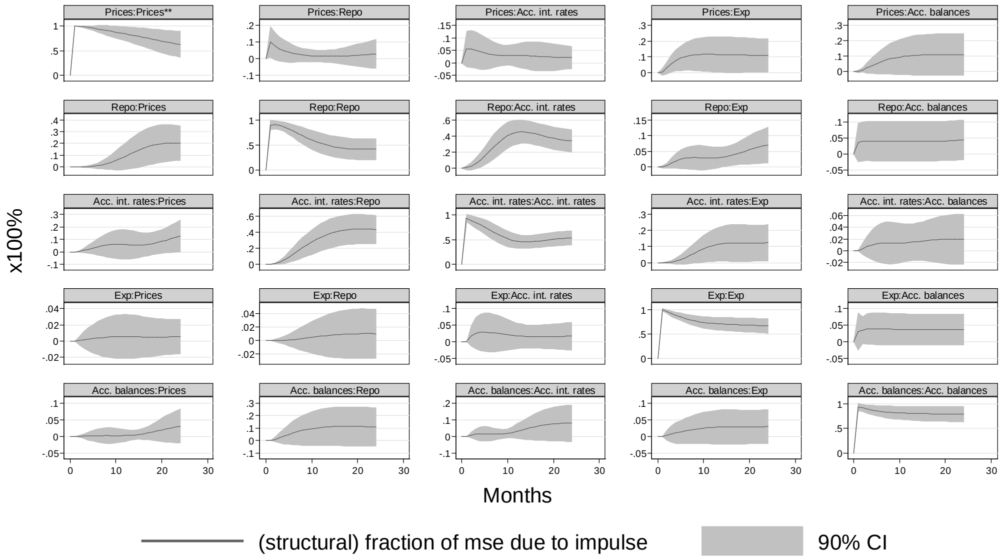
(i) Household cheque account IRFs (retail exp.; 1 lag)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

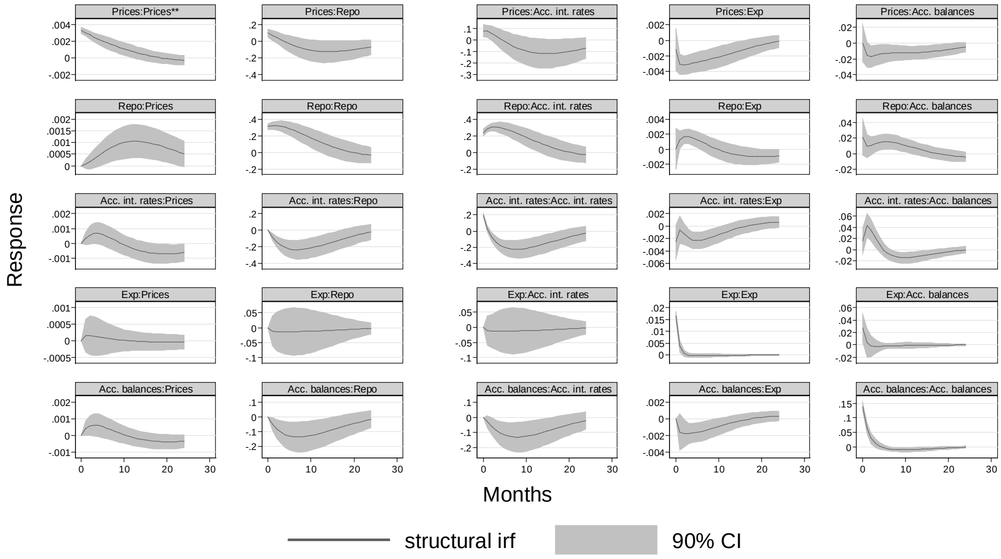
(ii) Forecast-error variance decomposition of household cheque account IRFs (retail expenditure)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

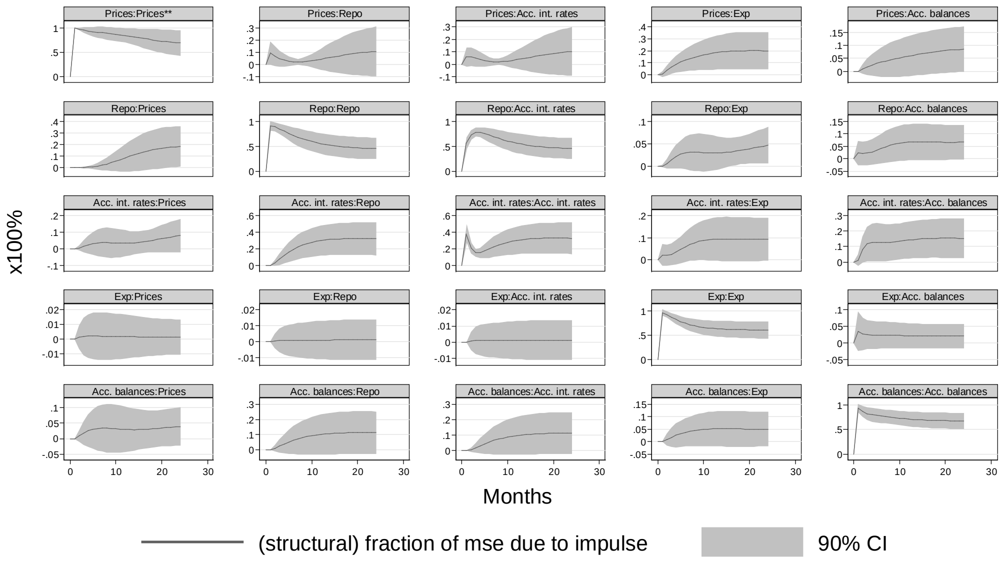
(iii) Household call account IRFs (retail exp.; 1 lag)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

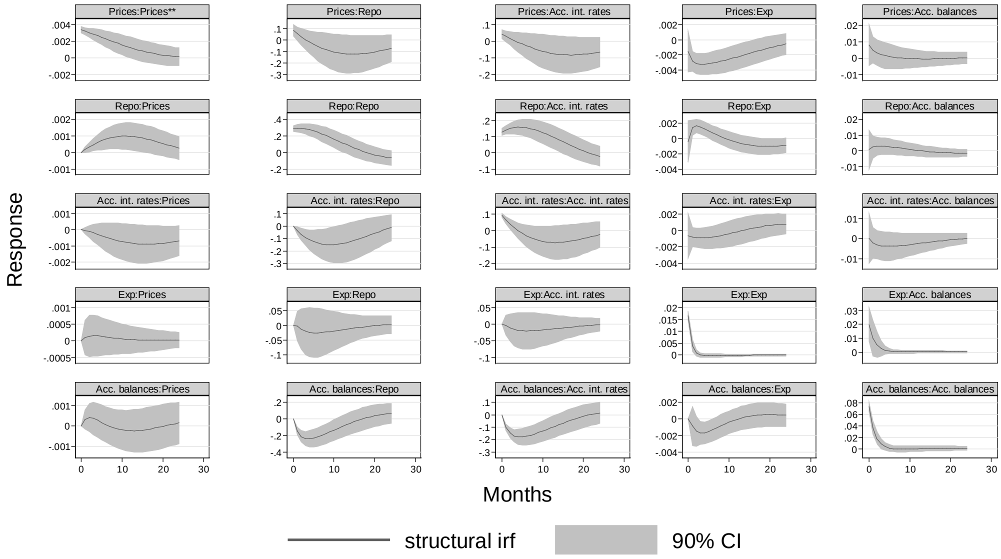
(iv) Forecast-error variance decomposition of household call account IRFs (retail expenditure)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

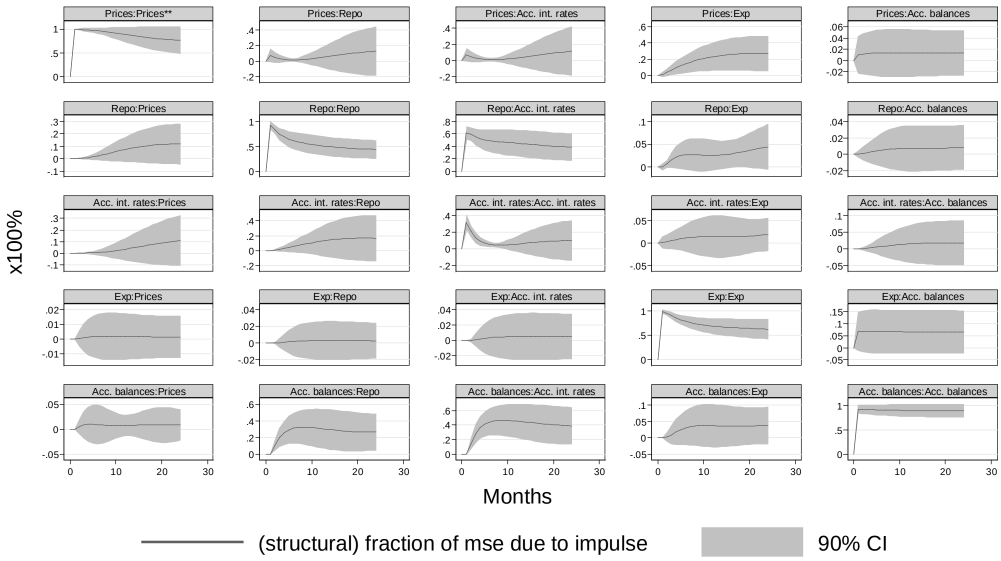
(v) Household savings account IRFs (retail exp.; 1 lag)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

(vi) Forecast-error variance decomposition of household savings account IRFs (retail expenditure)*



Notes:

* Response to a single std. dev. shock
** Impulse:Response

Chapter 4

Monetary Intermediation in a General Equilibrium Framework

"The day on which it suits the British merchant to purchase and send away a large quantity of goods, may not be that on which he finds it convenient to pay for them. If it is made necessary for him to give ready money in return, he must always have in his hands a very large stock of money; and for the expence [sic] of keeping this fund (an expence [sic] consisting chiefly in the loss of interest) he must be repaid in the price of the commodities in which he deals. He avoids this charge, and also obtains time for preparing and adjusting his pecuniary concerns, by buying on credit; that is to say, by paying for his goods not by money, but by the delivery of a note in which he promises the money on a future day"

— Thornton (1965 [1802], p. 76).

4.1. Introduction

Does monetary intermediation affect macroeconomic aggregates in a general equilibrium framework? If so, does this have implications for monetary policy? To investigate these questions, I extend the seminal monetary DSGE model of Cooley & Hansen (1989) to incorporate monetary intermediation by banks and an inflation targeting monetary authority.

Cooley & Hansen's model makes the important contribution of demonstrating that money supply dynamics affect real economic activity, at least in the short run. Despite this important result, the model has a number of empirical limitations. For example, unlike the characteristics of US data, the model does not exhibit a positive aggregate correlation between output and inflation, nor does it exhibit a positive liquidity effect. Money supply in the model is exogenous and is governed by a Friedman-style growth rule. By contrast, money supply today is a largely endogenous process driven by bank intermediation, and monetary policy is an endogenous inflation targeting process.

In this chapter, I extend the model of Cooley & Hansen to incorporate both endogenous bank intermediated money supply, and endogenous interest-rate-adjusting inflation targeting. Using this augmented model, I explore how monetary intermediation impacts on aggregate macroeconomic variables in a general equilibrium setting, and what this means for policy makers. The results indicate that monetary intermediation matters for the real economy. By comparison to the original Cooley & Hansen model, both monetary intermediation and inflation targeting monetary policy are shown to lower the standard deviations of real output, consumption, and investment in response to total factor productivity (TFP) shocks. This is due to consumption smoothing effects arising from both monetary intermediation services and inflation targeting interest rate adjustments. In addition, credit shocks increase the variance of real variables, but inflation targeting interest rate adjustment responses lower this variance. These results have interesting monetary policy implications that contribute to current debates. The results suggest that an interest-rate-adjusting inflation-targeting rule reduces aggregate real economic volatility and is a useful tool for reducing the negative impact of credit shocks in an endogenous monetary system. One caveat to these results is that the system is closed, and so open economy effects are not taken into account.[49] The results also address two fundamental *empirical* shortcomings within the original Cooley & Hansen model. First, positive credit shocks to the monetary intermediation process lower nominal interest rates and lead to an increase in money supply, output, and investment, generating a positive liquidity effect. The original Cooley & Hansen model does not generate a positive liquidity effect. Second, interest-rate-adjusting inflation targeting monetary policy leads to a positive aggregate correlation between output and inflation in the model. The original Cooley & Hansen model does not generate a positive correlation between output and inflation.

In choosing a flexible-price RBC model, as opposed to the more popular New Keynesian (NK) framework, I am not taking a position on the necessity of NK

[49] It would be interesting to extend the model structure to include the impact of open market effects for comparison, but this is left for future research.

frictions.[50] Instead, my focus is on the real substitution effects that follow due to the transformation of debt into money by banks. These substitution effects occur due to changes to *intertemporal* relative prices. I am interested in the *timing* friction that separates the exchange of G&S when G&S are exchanged using debt. I am interested in the macroeconomic impact of the mitigation of this timing friction by banks as a result of monetary intermediation. I am not looking at the distortions caused by pricing frictions and nominal rigidities.[51] Because I am interested in the effects of this time wedge and the change in intertemporal relative prices, and not on the effects of nominal frictions, a flexible-price model is a suitable framework. I use a CIA setup because it provides a useful model structure for imposing a timing friction that separates the exchange of G&S across time periods.

4.2. Context

4.2.1. Flexible-price models

Cooley & Hansen (1989) provide a seminal contribution on the real effects of money supply in the economy. They specify a flexible-price CIA model that is subject to stochastic shocks to an exogenous money growth policy rule. They illustrate that money can have important *real* effects, even in an environment where there are no nominal rigidities. This represented a direct challenge to the "classical dichotomy", which holds

---

[50] The New Monetarist models of Williamson & Wright (2010; 2011) and the I-theory of Brunnermeier & Sannikov (2016) are recent examples of other studies that also explore the role of money and banks in flexible-price environments.

[51] It is, however, worth emphasising that both a liquidity effect and a positive correlation between output and inflation are achieved without imposing nominal frictions.

that money is neutral and affects only nominal and not real variables. The model does not exhibit the neoclassical dichotomy because money effectively imposes an "inflation tax". This means that as anticipated inflation increases, money devalues and so agents hold less money. As a result, inflation effectively acts as a tax on G&S purchased with money, and essentially subsidises leisure (as leisure does not require money in order to be "consumed"). This impacts on household choices between labour (which earns the money wages required for G&S purchases) and leisure, with corresponding real effects. The shortcomings of the model, however, are that it does not accurately account for the empirical dynamics of the nominal variables, nor the interactions of various real and nominal variables. For example, the nominal interest rate has close to zero correlation with output in the model, and the model does not generate a positive correlation between output and inflation or between output and money growth.

In this chapter, I augment the model of Cooley & Hansen with both monetary intermediation and inflation targeting monetary policy. The specification of a bank's monetary intermediation technology and its balance sheet structure are derived from Gillman & Kejak (2011). Gillman & Kejak (2011) draw on a succession of earlier versions of a core model first developed by Gillman (1993). Gillman (1993) extends the standard stochastic CIA model by including a credit production function. Credit provides an alternative medium-of-exchange payment option to cash. Households choose between a foregone interest opportunity cost when holding cash, and the real resource cost of using credit when deciding which payment method to use. In the model, the marginal rate of substitution between labour and leisure therefore depends not only on interest rates, but also on the cost of credit provision by banks. Gillman, et al. (1997), Gillman &

Kejak (2005), Benk, et al. (2005), Gillman & Kejak (2011), and Davies, et al. (2012) add various extensions to the original core model.[52] The model presented in this chapter is a further extension of this framework.

Dittmar, et al. (2005) specify a flexible-price shopping-time model augmented with an inflation targeting central bank following a Taylor rule. The nominal interest rate is subject to stochastic policy shocks. They frame their discussion with the empirical observation that, whereas inflation and output are found to be highly persistent in US data with significant positive cross-correlations between inflation and output, this relationship is absent from the seminal CIA models. Dittmar, et al. explain that this empirical relationship has been broadly interpreted as evidence that monetary policy has large real effects. Dittmar, et al. however, demonstrate that this relationship can be accounted for in a flexible-price DSGE model if the central bank follows an interest rate rule, even if monetary policy shocks have *almost no real effects*. They investigate the effects of alternative monetary policy rules and find that inflation dynamics are determined by the dynamics of the spread between the real interest rate

[52] Gillman (1993) focuses on the welfare effects of inflation in this setting. Gillman finds that the change in the margins of substitution increases the welfare costs of inflation, and that the interest elasticity of money demand is more negative than in cash-only models. Gillman, et al., (1997) use the same framework to describe the income velocity of money, and show that financial innovation that reduces the cost of credit will reduce the proportion of transactions financed with cash. This increases the velocity of cash without shifting the money demand function. Gillman & Kejak (2005) add both capital and human capital to the model of Gillman (1993), and again focus on the welfare effects of inflation. Inflation in the model acts as a tax on both G&S produced and labour productivity, causing the real interest rate and the output growth rate to fall. This results in a negative non-linear inflation-growth relationship. Benk, et al. (2005) build on the model of the Gillman & Kejak (2005) by adding credit shocks. They show that a positive credit productivity shock increases consumption, output, and prices and decreases leisure, capital, and real money use. Gillman & Kejak (2011) extend Gillman & Kejak (2005) to include a decentralised banking sector and investment that is subject to the CIA constraint. Gillman & Kejak (2011) illustrate that in this setup, unlike standard CIA models, inflation reduces the investment rate, output growth, and the real interest rate. Davies, et al. (2012) also extend Gillman & Kejak (2005) to include a decentralised banking sector, but also include money that is supplied stochastically by the central bank. They derive a Taylor-rule-like general equilibrium condition, even though monetary policy is modelled in terms of a money growth rule.

and the central bank's interest rate target. Inflation persistence results from any shock that drives a persistent wedge in this spread.[53] They show that inflation persistence is sensitive to both the degree of interest rate smoothing and the central bank's reaction to output, with inflation autocorrelations being more sensitive to the weight on output than they are to the weight on inflation. They conclude that inflation persistence can therefore be expected to be pervasive in any economy where the central bank is using an interest rate rule.

Similar to Dittmar, et al. (2005), Chen (2008) uses a flexible-price shopping time model augmented with an inflation targeting Taylor rule to examine the impact of macroeconomic policy shocks. Chen notes that a standard flexible-price CIA monetary model with a simple money growth rule fails to generate a liquidity effect, persistence in output and inflation, or pro-cyclical inflation. By contrast, Chen's model generates (i) a liquidity effect; (ii) higher persistence in output and inflation, as well as positive unconditional cross-correlations between output and inflation; and (iii) less volatile consumption, hours worked, productivity, inflation, and nominal interest rates; but investment is more variable. Chen explains that a decrease in the nominal interest rate (expansionary monetary policy) raises money demand, decreases shopping time, and increases hours worked. The increase in labour supply increases the marginal productivity of capital, investment spending, and then output. This generates the liquidity effect. The real effects of a monetary shock are, however, small.

In this chapter, I add to this strand of modelling literature. I illustrate both how monetary intermediation by banks affects the behaviour of macroeconomic variables

[53] This has an interesting parallel to Wicksell's notion that price instability is driven by discrepancies between the "natural" (equilibrium) interest rate and the actual money rate of interest.

relative to established benchmarks, and I show how this endogenous money interacts with inflation targeting monetary policy (endogenous monetary policy) to influence business cycle characteristics. Like Dittmar, et al. (2005) and Chen (2008), I augment a standard RBC monetary model with a Taylor rule, but I do not use a flexible-price shopping time model. Instead, I specify a CIA RBC model and add both a Taylor rule, as well as a banking sector derived from Gillman & Kejak (2011). The role of banks in the model is to transform debt into deposits. Deposits provide a payment substitute to cash.

4.2.2. New Keynesian models

Bernanke, et al. (1999) provide an important contribution on the role of credit in determining the character of the business cycle in an NK DSGE framework. They specify a sticky price environment in which credit market frictions influence the transmission of monetary policy (there are, however, no formal banks in the model). Credit market frictions arise from asymmetric information between lenders and potential borrowers, and this results in a conflict of interest and a "costly state verification" agency problem. This leads to a difference between the cost of funds raised externally, and the cost of raising funds internally within firm. This difference is referred to as an "external finance premium". The external finance premium is assumed to depend inversely on borrower net worth; that is, when borrowers have little wealth to contribute to a project, the potential conflict of interest is greater, implying increased agency costs. These are costs for which lenders must be compensated with a higher premium. This setup generates a "financial accelerator" effect, meaning that: "to the extent that borrowers' net worth is pro-cyclical (because of the pro-cyclicality of profits and asset prices, for example), the external finance premium will be countercyclical, enhancing the swings in borrowing and thus in investment, spending, and production" (Bernanke, et al., 1999, p. 1345). Bernanke, et al. show that credit market frictions and fluctuations in borrowers' net worth may significantly amplify and propagate both real and nominal exogenous shocks to the economy through the financial accelerator mechanism.

Christensen & Dib (2008) estimate and simulate a sticky price financial accelerator model along the lines of Bernanke, et al. (1999), but wherein monetary policy is characterised by a Taylor rule. Their results show that the financial accelerator mechanism amplifies and propagates the impact of demand-side monetary policy and preference shocks on both investment and the price of capital. However, the financial accelerator mechanism dampens the rise of investment following positive supply-side technology and investment-efficiency shocks. They also find that the initial impact of the accelerator on output and inflation is relatively minor and that this is partly due to the response of monetary policy.

Jakab & Kumhof (2015) describe a series of NK DSGE models wherein access to bank intermediated credit provides agents with cash-like deposits (i.e., an endogenous money supply similar to the role of banks in this chapter). The models of Jakab & Kumhof are differentiated by the degree to which bank-lending-spread frictions inhibit the provision of bank intermediated credit. In response to credit shocks, the models with less price rigidity predict changes in the size of bank balance sheets that depend less on changes in lending spreads, are larger, happen faster, and have more pronounced real economic effects (i.e., more quantity and fewer price effects). These models also suggest there is more pro-cyclical bank leverage when lending-spread frictions are reduced.

In this chapter, I explore how endogenous money and endogenous monetary policy affect the features of the business cycle in a DSGE framework. Prices are flexible and bank intermediation transforms debt into a cash substitute. This model shows that monetary intermediation by banks amplifies the inflationary effects of exogenous money shocks of the kind modelled by Bernanke, et al. (1999), but this is not the case for real TFP shocks. Credit shocks are shown to be pro-cyclical, but inflation targeting monetary policy reduces the magnitude of the impact on real variables, much like in Christensen & Dib (2008). And like the models in Jakab & Kumhof (2015), the model presented in this chapter predicts pro-cyclical bank intermediation. The model therefore complements the NK literature, albeit in a flexible-price environment.

4.3. Model overview

Cooley & Hansen (1989) specify an archetypal CIA model, which they extend to incorporate a money growth rule subject to stochastic shocks. I further augment this CIA model in three ways. First, I broaden the definition of money to include not only government issued cash (i.e., outside or exogenous money) but also inside (or endogenous) bank-money. Banks are introduced into the model, and their purpose is to transform debt into money in the form of bank deposits. Second, investment spending is financed by household savings, which are channelled to firms via financial intermediaries (investment is therefore also subject to the CIA constraint). And third, I explore the effects of both exogenous monetary policy rules (in the form of money growth policy rules) and endogenous (interest-rate-managed inflation targeting) monetary policy rules by comparing results from models with these different rule types.

Banks play the role of monetary intermediaries within the model. They transform debt into money. I introduce monetary intermediation into the model by exploiting the fact that any expenditure subject to a CIA constraint represents expenditure financed by endowments *earned in a previous period*, but which are only *paid over to households in the current period*. The CIA constraint can therefore be thought of as money that is tied up in some form of debt-credit contract. These contracts *delay the purchasing power* of agents who are unable to spend money that is locked into these contracts until they mature. The intuition is that, in practise, many transactions take place on the basis of debt-credit agreements. For example, wages are often paid only at the end of some working period (the end of a week or the end of a month for example); bonuses and profit shares are only distributed at the end of some accounting period (the end of a month or the end of a year for example); and firms often use working capital arrangements that delay payment to suppliers for a specified time period (30, 60, or 90 days as examples).[54] The CIA constraint is a tool for introducing these payment frictions into the model. The CIA constraint in this model is therefore not interpreted as shorthand for introducing money demand *per se* (money demand is the usual motivation for a CIA constraint), rather, the CIA constraint is used here as an extension

[54] This can be thought of as a formal method of modelling Schumpeter's "synchronisation" argument for holding money (Schumpeter, 1939, p. 547). In effect, what is happening here, is that the implicit assumption in general equilibrium models—that every agent buys and sells G&S simultaneously—is relaxed. And rather than arbitrarily separating who transacts first—as in Sissoko (2007)—economic rationale determines that production and employment precede wage income and consumption.

of the RBC model to include a formal timing separation (i.e., a temporal friction) between the *purchase* and the *payment* components of many exchange transactions.[55]

I assume all exchange in the model is conducted using money (i.e., there is a transactions demand for money). However, some exchanges are subject to debt-credit agreements that delay purchasing power for those agents who extend credit to others. These agreements conform to a standard CIA constraint, but with a different underlying interpretation of the constraint. The structure of the model is as follows:

- cash and bank deposits offer substitute means of payment, but ultimately, all cash payments are deposited into bank deposit accounts;
- money therefore circulates in the form of bank deposit balances; within each period, firms accumulate money from the sale of G&S;
- money is then transferred to households in payment of wages and profits, but only at the end of each period;
- households then spend this money, buying either financial assets or G&S, but only from the beginning of the following period;
- paying wages in this way represents interest-free credit to firms, which is implicitly provided by households;
- waiting for wages and profits to be paid constrains household expenditure;
- this constraint can be relaxed (as is common in practise) through access to bank loans.

[55] This temporal separation follows the same intuition as the setups found in the New Monetarist literature discussed in the literature reviewed in Chapter 1.

Banks, therefore, essentially enable households to bring their purchasing power forward in time. Banks are mechanisms that increase liquidity by relaxing the households' consumption timing constraint. They do so by transforming debt into money in the form of bank deposits. The *raison d'être* of banks in the model is to provide these monetary intermediation services.

I specify a variety of model specifications according to different monetary policy rules. I consider both a cash growth rule corresponding to that of Cooley & Hansen (1989)—Model v.1—and various Taylor (1993) rule specifications—Model v.2. By combining money supply and policy rule features, I am able to explore how both the form of money and the form of monetary policy rule impact on empirical relationships observed among macroeconomic variables within the models. This is a direct continuation of the work begun by Cooley & Hansen (1989). Comparison is made to benchmark results from the literature.

4.4. The model

I modify a standard, flexible-price CIA DSGE model to include both monetary intermediation by banks and inflation targeting monetary policy. The agents populating the economy include households, firms, banks, asset managers, and government. Agents treat bank deposits as money, and households have access to bank intermediated credit. Firm investment expenditure is financed by borrowing money from households, and money is channelled from households to firms via financial markets. Agents take prices as given.[56]

Following the literature, I make specific assumptions about the functional form of the preferences and constraints faced by the representative agents. These standard assumptions take the form of incremental deviations from the well-understood neoclassical growth model, and are mostly constrained by a need for tractability—as discussed by Karagedikli & Matheson (2010, p. 116). The solution form is that of a social planner maximising the expected utility of the representative individual by picking a path for consumption and leisure subject to constraints. Household and firm parameter calibrations are taken from Walsh (2010). The calibration of the banking sector technology is taken from Benk, et al. (2005). Inflation targeting is implemented in the form of a forward looking Taylor rule, as per Taylor (1993). Various Taylor rule calibrations are considered, compared, and discussed. Asset markets open and monetary shocks occur *prior* to G&S markets opening.[57]

4.4.1. Households

The perfectly competitive economy is populated by a large number of infinitely lived households with identical preferences. Households participate in the G&S, capital, financial, and labour markets.

[56] I adapt and combine elements from the models of Stockman (1981), Schorfheide (2014b), Benk, et al., (2005), Scheffel (2010), Walsh (2010), and Gillman & Kejak (2011) for this purpose.

[57] See Salyer (1991) and Walsh (2010, pp. 101, 210) for detailed discussions on the technical implications of these timing conventions.

4.4.1.1. Household utility

The representative household seeks to maximise utility over time, such that:

$$\max_{c_t, q_t} E_0 \sum_{t=0}^{\infty} (1+\zeta)^{-t} u(c_t, q_t)$$

where c_t and q_t are household consumption and leisure respectively in time period t, and parameter ζ is the (constant) subjective rate of discount.[58] The expectations operator, E_0, indicates that the expectations are formed conditional on information available in the current period. Available information is assumed to consist of all current and lagged values of all variables, but future values are unknown. The utility function, $u(\cdot)$, is assumed to be strictly concave and twice continuously differentiable, such that marginal utility is positive $(u_c, u_q > 0)$ and diminishing $(u_{cc}, u_{qq} < 0)$. Additionally, the Inada (1963) conditions are assumed to hold such that $\lim_{c,q \to 0} u_c, u_q = \infty$ and $\lim_{c,q \to \infty} u_c, u_q = 0$. This ensures an interior solution (Lucas, 1987b). The explicit functional form of the utility function is a combined adaptation of Benk, et al. (2005) and Walsh (2010), such that:

$$\max_{c_t, h_t, n_t} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{{c_t}^{1-\phi}}{1-\phi} + \psi \frac{(1-h_t-n_t)^{1-\eta}}{1-\eta} \right], \qquad 0 < \beta < 1\,, \qquad 1 \leq \psi, \eta, \phi \qquad (4.1)$$

where $\beta = (1+\zeta)^{-1}$ is the subjective discount factor, ψ is the preference parameter for leisure, h_t represents real labour resource costs of bank credit production at time t, and n_t is labour hours supplied to G&S production at time t. The household time endowment is normalised to 1 so that $1 - h_t - n_t$ represents leisure. At time t, the

[58] The subjective rate of discount is also called the rate of time preference in the literature.

representative household chooses to allocate its time endowment between leisure and working hours, with $u_n, u_h < 0$ and $u_{nn}, u_{hh} > 0$. Household utility is inter-temporally additive and separable, and is separable in its arguments (i.e., intra-temporally separable). This has the advantage of isolating the effects of money on consumption from those on labour hours. Utility is a power function in both consumption and leisure, and exhibits constant relative risk aversion (CRRA).[59] CRRA ensures $u_{cn}, u_{ch} \geq 0$, and therefore permits a steady state when combined with a standard Cobb-Douglas production function (King, et al., 1988). The coefficient of relative risk aversion is given by ϕ, and therefore ϕ^{-1} is the intertemporal elasticity of substitution (IES) of consumption with respect to the real interest rate. The IES of consumption is a measure of the curvature of the utility function, and determines the speed of adjustment to the steady-state equilibrium; it is therefore responsible for consumption smoothing effects. It contributes to the basic impulse response mechanism within the RBC framework by determining the amplitude of the initial response to a shock, and by influencing the characteristics of its propagation.[60] The IES of labour is given by the expression η^{-1}.[61] The higher the IES of labour, the more households work if wages increase, and therefore, the more households shift the path of labour supply in response to temporary changes in the wage rate. The labour supply response to changes in wages is therefore

[59] The use of power utility is convenient because "it nests two popular special cases in the real business cycle literature: log-utility for leisure in a model with divisible labor and linear derived utility for leisure in a model with indivisible labor" (Campbell, 1994, p. 482). The former case corresponds to $\eta = 1$, the latter to $\eta = 0$. Christiano & Eichenbaum (1992) and King et al., (1988) provide a comparison of these two cases.

[60] In response to a positive AR(1) TFP shock, a lower ϕ results in a relatively lower initial jump in consumption and more of a hump shaped (less smooth) consumption impulse response. It also results in a higher initial jump in output, investment, and hours worked, and a lower jump in real wages (Walsh, 2010). As stated by Cooley & Prescott (1995, p. 17): "variations in the intertemporal elasticity of substitution affect transitions to balanced growth paths but not the paths themselves".

[61] The IES of labour in this setup is equal to the labour supply elasticity, also called the Frisch elasticity after Frisch (1967 [1933]).

an additional propagation mechanism. Households maximise utility subject to the following CIA and budget constraints.

4.4.1.2. The cash-in-advance constraint

The CIA constraint proposed by Clower (1967) and developed formally by Grandmont & Younes (1972) and Lucas (1980; 1982), explicitly requires that money is used to purchase G&S such that:

$$P_t c_t \leq M_{t-1}$$

where P_t is the general price level,[62] c_t is real household consumption, and M_{t-1} is the amount of cash held in advance of transacting.[63] Similarly, in the model setup in this chapter, consumption is funded by income, but I assume wages, profit share, and interest payments (the types of income earned by households) are only paid at the very end of the period they are accrued, and so can only be spent from the beginning of the following period. Therefore, I interpret M_{t-1} as the value of income-earning contracts entered into at the beginning of the previous period that are paid only at the very end of the previous period, and that are therefore available to be spent (as money) only from the beginning of the following period. Income is therefore subject to a delayed payment constraint—i.e., the CIA constraint.

[62] I adopt the convention of using lower case lettering to refer to real variables and upper case lettering for nominal variables.
[63] M_{t-1} is defined as aggregate nominal money demand in these models.

Real investment (x_t) in the model is also subject to a delayed payment constraint:[64] firms finance investment by borrowing money from households (through financial markets), which is only repaid at the end of the period. Firms receive nominal lump-sum cash transfers (taxes, if negative) from government (T_t), and households have access to nominal bank intermediated loans (F_t). Households and firms purchase consumption G&S and investment G&S using money in the form of deposits originating from new government cash transfers, or new bank loans. Therefore, in *real* terms,[65] the CIA constraint in the model takes the form:

$$c_t + x_t \leq \frac{m_{t-1}}{1 + \pi_t} + \tau_t + f_t.$$

where $P_t = P_{t-1}(1 + \pi_t)$; π_t is the rate of inflation; $m_t = \frac{M_t}{P_t}$, $\tau_t = \frac{T_t}{P_t}$; and $f_t = \frac{F_t}{P_t}$.

I make the simplifying assumption used by Benk, et al. (2005), Scheffel (2010), and Gillman & Kejak (2011) that all cash is deposited with banks and that all expenditure passes through the banking system. Household consumption is paid for using demand deposits which circulate within the banking system. All capital forming investment spending is channelled via asset managers, but funds channelled to firms by asset managers are ultimately spent through the banking system. Government transfers are paid directly into bank accounts and firms both receive revenue and pay household income via bank accounts. All money is therefore held within the banking system in the form of deposits. The banking system has a typical balance sheet structure: on the asset

[64] See Stockman (1981) and Abel (1985) as examples of models that include investment in the CIA constraint.
[65] From this point forward, equations are given in real terms only for convenience and brevity, and the current period general price level is implicitly normalised to unit value.

side are bank loans (f_t) and cash holdings $\left(\frac{m_{t-1}}{1+\pi_t}+\tau_t\right)$; [66]on the liabilities side are the bank deposits (d_t), which finance consumption (c_t) and investment (x_t) expenditures such that:

$$c_t + x_t = d_t.$$

The CIA constraint can therefore be written as a bank solvency constraint:

$$d_t \leq \frac{m_{t-1}}{1+\pi_t} + \tau_t + f_t.$$

The CIA constraint is therefore introduced into the model as a bank solvency constraint and not as an aggregate money demand function.

The bank balance sheet solvency constraint can be specified from a household perspective as:[67]

$$c_t + s_t \leq \frac{m_{t-1}}{1+\pi_t} + \left(1 + i_{s,t-1}\right)\frac{s_{t-1}}{1+\pi_t} + f_t, \qquad \forall\, t \geq 0 \tag{4.2}$$

where f_t is expenditure financed by bank-credit;[68] households begin each period with money paid out from previous period contracts, $\frac{m_{t-1}}{1+\pi_t}$ and money received from gross interest payouts of matured bond holdings, $\left(1 + i_{s,t-1}\right)\frac{s_{t-1}}{1+\pi_t}$; and households transfer money to firms for investment via new bond purchases through financial

[66] The banking system's current cash assets are the amount of cash injected by government at the beginning of this period and the cash value of the delayed payment contracts paid to households at the end of the previous period. The value of these contracts is deposited as cash within the banking system and is transferred to households (through a direct debit within the banking system for example) at the end of the previous period. It is therefore held as cash by the banks and is available for spending in the form of bank deposit balances by households at the beginning of the current period.
[67] Investment (x_t) and real lump-sum cash transfers from government (τ_t) are connected to households only indirectly via financial market purchases of single period savings bonds (s_t). See Sections 4.4.2.2 and 4.4.4.1 for further details of how household savings finance firm investments via financial markets.
[68] See Section 4.4.3 for the discussion on credit production.

intermediaries, s_t. The inclusion of s_t as an explicit argument within the household constraint indicates that asset markets open for trade prior to G&S markets—what Kohn (1984) refers to as a "liquid asset market". The household is therefore able to adjust its portfolio between cash and savings before entering the G&S market to purchase consumption G&S.[69] All transactions are frictionless, meaning there is costless conversion of savings into cash or deposits and *vice versa*.

4.4.1.3. The household budget constraint

The representative household receives income and interest from labour supplied at the equilibrium real wage rate, from dividend yielding equity (physical capital is assumed to be owned by firms and firms by households respectively), and from interest yielding savings. Wages are agreed (at the beginning of a period) in the form of non-interest bearing contracts that are paid only at the end of a period (and thus wages only yield monetary purchasing power from the beginning of the following period). Households purchase consumption G&S and savings products (bonds) which are paid for with money. The household budget constraint is therefore given by:

$$c_t + s_t + m_t \leq \frac{m_{t-1}}{1+\pi_t} + \left(1 + i_{s,t-1}\right)\frac{s_{t-1}}{1+\pi_t} + w_t n_t + \Pi_t^F + \Pi_t^B + \Pi_t^I, \qquad \forall\, t \geq 0 \qquad (4.3)$$

where w_t is the hourly real wage rate; $1 + i_{s,t-1}$ is gross nominal interest payments on savings invested in financial markets in the previous period; and $\Pi_t^F + \Pi_t^B + \Pi_t^I$ are real profit shares received from firms, banks, and asset managers respectively. As per

[69] This timing ensures that the CIA constraint is always binding as long as the nominal interest rate is positive (Salyer, 1991).

Woodford (2003), an implicit assumption is that financial markets are complete and that the "available financial assets completely span the relevant uncertainty faced by households about future income, prices, taste shocks, and so on, so that each household faces a single intertemporal budget constraint".

Comparing the CIA constraint Equation (4.2) with the budget constraint Equation (4.3) shows that wages and profits earned during period t are unavailable for financing consumption in period t. Transfers and payments for G&S sold are implicitly paid into firm bank accounts during a period, but payment of wages and distributed profits are only transferred to household bank accounts at the end of the period, and are therefore available for spending by households only at the beginning of the following period. The CIA constraint is traditionally interpreted as an exogenously imposed liquidity demand for money. However, in this model, I interpret the CIA constraint as a formal timing friction that constrains the spending ability of agents whose funds are tied up in debt-credit contracts. This delayed payment constraint can be relaxed by the monetary intermediation services provided by banks.

4.4.1.4. Household optimisation

The problem solved by the representative household is to maximise Equation (4.1) subject to a sequence of flow constraints given by Equations (4.2) and (4.3). This problem can be solved using the method of Lagrangian multipliers[70] in terms of which the first-order conditions (FOCs) for the representative household's decision problem

[70] An alternative is to use the Bellman equation.

must be satisfied along with the transversality conditions.[71] The system does not have an analytical solution, however, the solution can be approximated around the steady state by log-linearising the steady-state equations. With λ_t as the Lagrangian multiplier on the budget constraint (i.e., the shadow price of income) and μ_t as the multiplier on the CIA constraint (i.e., the implicit utility cost of debt-credit contracts), the Lagrangian optimisation function is given as:

$$\mathcal{L} = \max_{c_t, n_t, h_t, s_t, m_t} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \left[\frac{{c_t}^{1-\phi}}{1-\phi} + \psi \frac{(1 - h_t - n_t)^{1-\eta}}{1-\eta} \right] \right.$$

$$+\mu_t \left[\frac{m_{t-1}}{1+\pi_t} + \left(1 + i_{s,t-1}\right) \frac{s_{t-1}}{1+\pi_t} - s_t - c_t + f_t \right]$$

$$\left. +\lambda_t \left[\frac{m_{t-1}}{1+\pi_t} + \left(1 + i_{s,t-1}\right) \frac{s_{t-1}}{1+\pi_t} + w_t n_t + \Pi_t^F + \Pi_t^B + \Pi_t^I - s_t - c_t - m_t \right] \right\}.$$

The FOCs with respect to c_t, n_t, h_t, s_t, and m_t are:

$$\mathcal{L}_{c_t}: \beta^t \left\{ \left[{c_t}^{-\phi} \right] + \mu_t \left[-1 + f_{c_t} \right] + \lambda_t [-1] \right\} = 0$$

$$\Rightarrow {c_t}^{-\phi} = \lambda_t + \mu_t \left(1 - f_{c_t}\right) \tag{4.4}$$

$$\mathcal{L}_{n_t}: \beta^t \{ [(\psi)(-1)(1 - h_t - n_t)^{-\eta}] + \lambda_t w_t \} = 0$$

$$\Rightarrow \psi (1 - h_t - n_t)^{-\eta} = \lambda_t w_t \tag{4.5}$$

[71] The transversality conditions $\left(\lim_{t \to \infty} \beta^t u_{c_t} z_t = 0;\ z = \{k, s, m\}\right)$ essentially prevent savings from becoming too large and ensure a unique optimal path. The optimal solution, however, effectively ignores the transversality conditions, as solutions are approximated close to the steady state.

$$\mathcal{L}_{h_t}: \beta^t\{[(\psi)(-1)(1-h_t-n_t)^{-\eta}] + \mu_t[f_{h_t}]\} = 0$$

$$\Rightarrow \psi(1-h_t-n_t)^{-\eta} = \mu_t(f_{h_t}) \tag{4.6}$$

$$\mathcal{L}_{s_t}: \beta^t\{\mu_t[-1] + \lambda_t[-1]\} + \beta^{t+1}\left\{\mu_t\left[\frac{1+i_{s,t}}{1+\pi_{t+1}}\right] + \lambda_{t+1}\left[\frac{1+i_{s,t}}{1+\pi_{t+1}}\right]\right\} = 0$$

$$\Rightarrow \mu_t + \lambda_t = \beta E_t(\lambda_{t+1} + \mu_{t+1})\left(\frac{1+i_{s,t}}{1+\pi_{t+1}}\right) \tag{4.7}$$

$$\mathcal{L}_{m_t}: \beta^t\{\lambda_t[-1]\} + \beta^{t+1}\left\{\mu_{t+1}\left[\frac{1}{1+\pi_{t+1}}\right] + \lambda_{t+1}\left[\frac{1}{1+\pi_{t+1}}\right]\right\} = 0$$

$$\Rightarrow \lambda_t = \beta E_t\left(\frac{\lambda_{t+1} + \mu_{t+1}}{1+\pi_{t+1}}\right). \tag{4.8}$$

Equation (4.4) illustrates that on the optimal path, the marginal utility of consumption exceeds the marginal utility of income (λ_t). In other words, monetary intermediation by the banking system lowers the costs of having to wait for debt-credit contracts to mature before they can be spent. Banks therefore provide liquidity services by bringing forward the purchasing power of households. At equilibrium, the marginal utility of consumption equals the "price" of consumption, consisting of the marginal utility of income plus the marginal utility of delayed purchasing power contracts. Equation (4.4) illustrates that the effect of the CIA constraint is to impose an opportunity cost on consumption, but that this cost is reduced via bank intermediation. Equation (4.5) is the labour supply equation equating the marginal utility of leisure with the product of the wage rate and the marginal utility of income. Equation (4.6) is labour supply to monetary intermediation. Equation (4.7) is the inter-temporal asset pricing Euler equation for savings. Equation (4.8) is the inter-temporal asset pricing Euler equation for delayed payment contracts, and it reflects the timing constraint that income received in period t cannot be spent on consumption until period $t + 1$.

Combining Equations (4.7) and (4.8) expresses the nominal interest rate as the ratio of the Lagrangian multipliers $i_{s,t} = \frac{\mu_t}{\lambda_t}$, and illustrates that if the nominal interest rate is positive, the CIA constraint must be binding (i.e., $\mu_t > 0$). In other words, if the interest rate on savings is positive, the institutional structure must be such that the CIA constraint is binding, otherwise households would prefer to hold their wealth in interest yielding savings bonds rather than non-interest yielding debt-credit contracts. From Equations (4.4), (4.7), and (4.8), the equilibrium opportunity cost of these contracts is $i_{s,t}$, and $1 + i_{s,t}\left(1 - f_{c_t}\right)$ is the "price" of consumption expenditure (Walsh, 2010, p. 105). If $\mu_t = 0$ but $i_{s,t} > 0$, no debt-credit payment contracts would be held in equilibrium and Equation (4.8) would not be a necessary condition for a solution. But because households must hold these contracts, a positive nominal interest rate acts as a "tax" on consumption and raises the "price" of consumption above its "production cost".[72]

Combining Equations (4.4), (4.5), (4.7), and (4.8) yields the intra-temporal (i.e., the choice between labour and leisure) labour supply optimisation condition:

$$\frac{\psi(1 - h_t - n_t)^{-\eta}}{{c_t}^{-\phi}} = \frac{w_t}{1 + i_{s,t}\left(1 - f_{c_t}\right)} \tag{4.9}$$

where the ratio of marginal utility between leisure and consumption is equated to the ratio of their respective marginal costs. From Equation (4.9), the CIA constraint introduces a "wedge of inefficiency" (Lucas, 1987a), which affects real variables by

[72] In other words, if the price of output is normalised to unit value, the actual cost for a unit of output incurred by a household includes both the unit production cost price *and* the additional nominal interest rate foregone by holding non-interest paying delayed payment contracts as opposed to savings.

distorting the marginal rate of substitution between consumption and leisure; monetary intermediation reduces the effects of this wedge if $1 \geq f_{c_t} > 0$.

4.4.2. Firms

The perfectly competitive economy contains a large number of identical firms that produce the economy's G&S output; utilising as inputs the state technology, the existing capital stock, and labour employed. Firms are assumed to own the economy's existing capital stock. Firms sell output to generate revenue from which they pay their wage bills and repay loans. Additionally, they borrow money through financial markets in order to invest in additional capital stock.

4.4.2.1. Firm objective function

The representative firm aims to maximise profit—defined as the difference between revenues and production costs. Profit is paid out to shareholder households in the form of dividends, Π_t^F, as per Nason & Cogley (1994) and Schorfheide (2000). Because households value a unit of date t dividends in terms of the consumption it enables in period $t+1$, date t dividends are discounted by the date $t+1$ marginal utility of consumption, such that the representative firm's objective function is:

$$\max_{\Pi_t^f, k_{t+1}, n_t, l_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left[\frac{\Pi_t^F}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] \tag{4.10}$$

subject to:

$$\Pi_t^F \leq y_t - x_t - w_t n_t + l_t - \left(1 + i_{l,t}\right) l_t, \qquad \forall\, t \geq 0 \tag{4.11}$$

where y_t is real output produced by the firm, x_t is investment, n_t is the labour hours employed, k_t is the beginning-of-period capital stock,[73] and l_t is cash borrowed (as a result of corporate bonds issued) through financial markets at the beginning of period t and repaid at the end of the period so that households are able to spend period t interest returns at the beginning of period $t + 1$.[74]

4.4.2.2. Production and technology

The representative firm produces real output using a constant returns-to-scale Cobb-Douglas production function with $y_{kn} \geq 0$, and borrows from financial markets to finance investment in capital, which evolves according to a law of motion such that:

$$y_t = e^{z_t} k_t^{\alpha} {n_t}^{(1-\alpha)}, \qquad 0 < \alpha < 1$$

$$z_t = \rho_z z_{t-1} + \varepsilon_{z,t}, \qquad \rho_z > 0, \qquad \varepsilon_{z,t} \sim \mathcal{N}(0, \sigma_z)$$

$$x_t \leq l_t$$

$$x_t = k_{t+1} - (1 - \delta) k_t$$

where α is the elasticity of output with respect to capital (and also represents capital's share of output), δ is the constant real depreciation rate of capital stocks, and e^{z_t} is a

[73] Occasionally in the literature, beginning-of-period capital stock timing is also given as k_{t-1}.
[74] Here investment is subject to the CIA constraint because it is financed by household income.

technology factor representing TFP, which is subject to a stochastic shock process. The technology shock, z_t, is a Solow (1957) residual,[75] and a first-order Markov process, which evolves according to an AR(1) law of motion. $\varepsilon_{z,t}$ is a white noise random variable; that is, it is serially uncorrelated and normally distributed with zero mean and constant variance ${\sigma_z}^2$. The persistence of the stochastic shock to technology is captured by the AR coefficient ρ_z, and it is an important propagation mechanism within the DSGE framework. With constant returns to scale, the behaviour of an individual firm is representative of aggregate behaviour without loss of generality.

4.4.2.3. *Firm optimisation*

In each period, firms invest in capital formation, purchase labour, and sell output in perfectly competitive markets. The problem solved by the representative firm is to maximise Equation (4.10) subject to (4.11), such that the Lagrangian optimisation function is given as:

$$\mathcal{L} = \max_{\Pi_t^F, k_{t+1}, n_t, l_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left\{ \left[\frac{\Pi_t^F}{{c_{t+1}}^{\phi}(1+\pi_{t+1})} \right] \right.$$

$$+\Lambda_{F1,t}[l_t - (k_{t+1} - (1-\delta)k_t)]$$

$$\left. +\Lambda_{F2,t}\left[e^{z_t} k_t^{\alpha} {n_t}^{(1-\alpha)} - (k_{t+1} - (1-\delta)k_t) - w_t^n n_t + l_t - \left(1 + i_{l,t}\right) l_t - \Pi_t^F\right] \right\}.$$

[75] As described by Danthine (1997, pp. 138-139): “Solow residuals are a black box summarizing what we do not know in the relation between factors of production and aggregate output and/or what we do not consider as a priority to model given the state of knowledge and the question at hand. One element of modern business cycle theory is the acceptance that there are forces driving the economy that are beyond economists’ duty or ability to explain. Political and climatic shocks are examples of the former, technological progress an example of the latter.”

The FOCs with respect to Π_t^F, k_{t+1}, n_t, and l_t are:

$$\mathcal{L}_{\Pi_t^F}: \beta^{t+1}\left\{\left[\frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})}\right] + \Lambda_{F2,t}[-1]\right\} = 0 \;\Rightarrow \Lambda_{F2,t} = \frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \qquad (4.12)$$

$$\mathcal{L}_{n_t}: \beta^{t+1}\left\{\Lambda_{F2,t}\left[(1-\alpha)\frac{y_t}{n_t} - w_t\right]\right\} = 0 \Rightarrow w_t = (1-\alpha)\frac{y_t}{n_t} \qquad (4.13)$$

$$\mathcal{L}_{k_{t+1}}: \beta^{t+1}\{\Lambda_{F1,t}[-1] + \Lambda_{F2,t}[-1]\}$$

$$+ \beta^{t+2}\left\{\Lambda_{F1,t+1}[(1-\delta)] + \Lambda_{F2,t+1}\left[\alpha\frac{y_{t+1}}{k_{t+1}} + (1-\delta)\right]\right\} = 0$$

$$\Rightarrow \Lambda_{F1,t} + \Lambda_{F2,t} = \beta E_t \Lambda_{F1,t+1}(1-\delta) + \beta E_t \Lambda_{F2,t+1}\left(\alpha\frac{y_{t+1}}{k_{t+1}} + 1 - \delta\right) \qquad (4.14)$$

$$\mathcal{L}_{l_t}: \beta^{t}\{[0] + \Lambda_{F1,t}[1] + \Lambda_{F2,t}[-i_{l,t}]\} = 0 \Rightarrow \Lambda_{F2,t}(i_{l,t}) = \Lambda_{F1,t}. \qquad (4.15)$$

The assumption of perfect markets implies that the representative firm makes zero profits in equilibrium, factor costs exhaust output, and both marginal capital and labour costs are equated to their respective marginal returns. Equation (4.13) is the labour demand function equating wages to the marginal product of labour. Combining Equations (4.12), (4.14), and (4.15) yields the inter-temporal asset pricing optimisation Euler condition (representing the household choice between consumption and saving), which states that the marginal utility of income is set equal to the present discounted value of the marginal utility of carrying income forward one period. In other words, it is the present discounted marginal benefit of foregoing consumption today in exchange for savings in the form of capital. Equations (4.4), (4.7), (4.8), (4.12), and (4.14) yield a modified[76] version of the Fischer equation:

$$1 + i_{l,t}\left(1 - f_{c_t}\right)\frac{1 + i_{l,t+1}}{1 + i_{l,t+1}\left(1 - f_{c_t}\right)} = (1 + r_t)(1 + \pi_t) + i_{l,t}(1 - \delta)(1 + \pi_t) \qquad (4.16)$$

where r_t is the real rate of return on capital (after depreciation). The standard Fischer equation defines gross nominal interest rates as the product of gross real returns on capital and inflation. Equation (4.16) indicates that a rising price level ($\pi_t > 0$) increases the opportunity cost of holding non-interest debt-credit contracts. From Equations (4.9) and (4.16), the CIA constraint leads to monetary frictions through an "inflation tax", resulting in substitution towards leisure and monetary intermediation provision. Positive inflation "taxes" non-interest debt-credit contracts, and therefore results in households substituting away from activities that rely on non-interest debt-credit contracts towards those that do not (i.e., away from the labour of G&S production, and towards leisure and bank intermediated credit provision). In other words, inflation redistributes in favour of borrowers.

4.4.3. Banks

The perfectly competitive economy contains a large number of identical banks. Banks provide households with monetary intermediation services by converting debt into bank deposits that provide a payment alternative to cash. Monetary intermediation is a

[76] Without bank intermediation, the left-hand side of Equation (4.16) would simplify to $1 + i_t$, and removing investment from the CIA constraint would remove the second term on the right-hand side, resulting in the standard Fischer equation, $1 + i_t = (1 + r_t)(1 + \pi_t)$.

function of credit technology, demand deposits (held for financing consumption),[77] and a real resource cost proxied by household labour.

4.4.3.1. Bank objective function

Like the representative firm, the representative bank aims to maximise profit, which is paid to shareholder households in the form of dividends (Π_t^B), and is discounted by the date $t+1$ marginal utility of consumption, such that the representative bank's objective function is:

$$\max_{\Pi_t^B, h_t, c_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left[\frac{\Pi_t^B}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] \tag{4.17}$$

subject to:

$$\Pi_t^B \leq f_t - f_{h_t} h_t - f_{c_t} c_t, \qquad \forall\, t \geq 0 \tag{4.18}$$

where f_t is the provision of bank intermediated credit, h_t is real resources in the form of labour hours directed towards credit provision, and c_t is bank demand deposits held at banks.[78]

[77] Time deposits finance investment through financial intermediation. Time deposits (unlike demand deposits) therefore are not part of the monetary intermediation function.

[78] Equation 4.18 is a profit maximising specification of objective function Equation 2.10 which includes the operating costs of monetary intermediation, but excludes the explicit costs of risk.

4.4.3.2. *Credit production and technology*

The representative bank extends loans in terms of a constant returns-to-scale Cobb-Douglas credit production function with $f_{ch} \geq 0$, such that:

$$f_t = \xi e^{v_t} h_t^{\gamma} {c_t}^{(1-\gamma)}, \qquad 0 < \gamma < 1$$

$$v_t = \rho_v v_{t-1} + \varepsilon_{v,t}, \qquad \rho_v > 0, \qquad \varepsilon_{v,t} \sim \mathcal{N}(0, \sigma_v)$$

where γ represents labour's share of credit output, ξ is a scale factor, and e^{v_t} is the bank's credit technology factor, which is subject to a stochastic shock process. Banks face increasing real marginal costs to credit provision. This enables the model to be solved for a unique equilibrium involving the use of both cash and bank-money (Gillman & Kejak, 2011). With constant returns to scale, the behaviour of an individual bank is representative of aggregate behaviour without loss of generality.

4.4.3.3. *Bank optimisation*

In each period, banks make use of both real labour resources and household willingness to use bank demand deposits as a medium-of-exchange alternative to cash in order to provide monetary intermediation services in perfectly competitive markets. The problem solved by the representative bank is to maximise Equation (4.17) subject to Equation (4.18), such that the Lagrangian optimisation function is given as:

$$\mathcal{L} = \max_{\Pi_t^B, h_t, c_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left\{ \left[\frac{\Pi_t^B}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] \right.$$

$$\left. + \Lambda_{B,t}\left[\xi e^{v_t} h_t^{\gamma} c_t{}^{(1-\gamma)} - f_{h_t} h_t - f_{c_t} c_t - \Pi_t^B\right] \right\}.$$

The FOCs with respect to Π_t^B, h_t, and c_t are:

$$\mathcal{L}_{\Pi_t^B}\colon \beta^{t+1}\left\{\left[\frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})}\right] + \Lambda_{B,t}[-1]\right\} = 0 \Rightarrow \Lambda_{B,t} = \frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \tag{4.19}$$

$$\mathcal{L}_{h_t}\colon \beta^{t+1}\left\{\Lambda_{B,t}\left[\gamma \xi e^{v_t} h_t^{\gamma-1} c_t{}^{(1-\gamma)} - f_{h_t}\right]\right\} = 0$$

$$\Rightarrow f_{h_t} = \gamma \xi e^{v_t} h_t^{\gamma-1} c_t{}^{(1-\gamma)} = \gamma \frac{f_t}{h_t} \tag{4.20}$$

$$\mathcal{L}_{c_t}\colon \beta^{t+1}\left\{\Lambda_{B,t}\left[(1-\gamma)\xi e^{v_t} h_t^{\gamma} c_t{}^{(-\gamma)} - f_{c_t}\right]\right\} = 0$$

$$\Rightarrow f_{c_t} = (1-\gamma)\xi e^{v_t} h_t^{\gamma} c_t{}^{(-\gamma)} = (1-\gamma)\frac{f_t}{c_t}. \tag{4.21}$$

The assumption of perfect markets implies that the representative bank makes zero profits in equilibrium, and input factor costs exhaust credit output and are equal to their respective marginal returns. Equation (4.20) is the bank's labour resource demand function and Equation (4.21) is the household deposit demand function. Combining Equations (4.5), (4.6), (4.13), and (4.20) yields:

$$w_t = (1-\alpha)\frac{y_t}{n_t} = \gamma \frac{f_t}{h_t}\frac{\mu_t}{\lambda_t}. \tag{4.22}$$

At the competitive equilibrium, labour is paid its marginal revenue product, and equation (4.22) equates labour's marginal revenue product across G&S and credit production. Equation (4.7) therefore indicates that the "price" of bank intermediated credit is $\frac{\mu_t}{\lambda_t}$, and from Equations (4.7) and (4.8), $\frac{\mu_t}{\lambda_t} = i_{s,t}$. Therefore, along the optimal path, the price of monetary intermediation is set equal to the opportunity cost of holding a CIA contract.

4.4.4. *Asset managers*

The perfectly competitive economy contains a large number of identical asset managers that channel funds within financial markets from saver households to borrower firms.

4.4.4.1. *Asset manager objective function*

Like the representative firm and the representative bank, the representative asset manager aims to maximise profit, which is paid to shareholder households in the form of dividends (Π_t^I), and is discounted by the date $t+1$ marginal utility of consumption, such that the representative asset manager's objective function is:

$$\max_{\Pi_t^B, s_t, l_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left[\frac{\Pi_t^I}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] \tag{4.23}$$

subject to:

$$\Pi_t^I \leq \left(1 + i_{l,t}\right) l_t - l_t - \left(1 + i_{s,t}\right) s + s_t + \tau_t, \qquad \forall\, t \geq 0 \tag{4.24}$$

where l_t is money loans made to firms through corporate bonds for the purposes of investment; s_t is household money lent to firms by purchasing bonds through asset managers at the beginning of period t, and is repaid at the end of the period such that households receive period t interest returns at the beginning of period $t+1$; and τ_t is government transfers.[79] All government transfers are channelled directly to firms like the setup in Nason & Cogley (1994) and Schorfheide (2000), such that:

$$l_t \leq \tau_t + s_t$$

4.4.4.2. *Asset manager optimisation*

Without any financial frictions, the asset manager solves the trivial problem:

$$\mathcal{L} = \max_{\Pi_t^I, s_t, l_t} E_0 \sum_{t=0}^{\infty} \beta^{t+1} \left\{ \left[\frac{\Pi_t^I}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] + \Lambda_{I1,t}[s_t + \tau_t - l_t] \right.$$
$$\left. + \Lambda_{I2,t}\left[\left(1 + i_{l,t}\right)l_t - l_t + s_t - \left(1 + i_{s,t}\right)s_t + \tau_t - \Pi_t^I\right] \right\}.$$

The FOCs with respect to Π_t^I, s_t, and l_t are:

$$\mathcal{L}_{\Pi_t^I} : \beta^{t+1} \left\{ \left[\frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \right] + \Lambda_{I2,t}[-1] \right\} = 0 \Rightarrow \Lambda_{I2,t} = \frac{1}{c_{t+1}{}^{\phi}(1+\pi_{t+1})} \tag{4.25}$$

$$\mathcal{L}_{l_t}, \mathcal{L}_{s_t} : \Lambda_{I2,t}(i_t) = \Lambda_{I1,t}. \tag{4.26}$$

From Equation (4.26), the equilibrium nominal interest rate charged on loans is equal to the nominal interest rate paid on savings.

[79] Government cash injections therefore enter the household CIA constraint via loans to firms.

4.4.5. Government

The specification of government's role is taken from Cooley & Hansen (1995). The government operates as monetary and fiscal authority subject to the following budget constraint:

$$g_t + \tau_t + \left(1 + i_{s,t-1}\right)\frac{b_{t-1}}{1+\pi_t} = m_t - \frac{m_{t-1}}{1+\pi_t} + b_t, \qquad \forall\, t \geq 0$$

where g_t is real government consumption and b_t is the value of government issued single period bonds.

4.4.5.1. Stochastic money growth (Model v.1)

Without any explicit inflation targeting monetary policy rule, money stocks are assumed to be realisations of an exogenous stochastic process. Of interest is the impact of monetary shocks and not the impact of changes in government spending. Therefore, g_t is set constant and equal to zero. With no loss in generality, b_t is also set equal to zero, implying that no government bonds are issued in the steady state.[80] The model therefore effectively ignores public finance considerations and the government budget constraint simplifies to:

$$\tau_t = m_t - \frac{m_{t-1}}{1+\pi_t}. \tag{4.27}$$

Letting the stochastic expression e^{u_t} denote the gross growth rate of money supply, then the government budget constraint is given by:

[80] See Cooley & Hansen (1995, p. 197) for a detailed explanation.

$$m_t = e^{u_t} \frac{m_{t-1}}{(1+\pi_t)}, \qquad \forall\, t \geq 0. \tag{4.28}$$

Monetary policy in this setup is therefore exogenously determined, and real lump-sum transfers amount to $(e^{u_t} - 1)\frac{m_{t-1}}{(1+\pi_t)}$. The random variable u_t is a first-order Markov process assumed to evolve according to an AR(1) law of motion:

$$u_t = \rho_u u_{t-1} + \varepsilon_{u,t}, \qquad |\rho_u| < 1, \qquad \varepsilon_{u,t} \sim N(0, \sigma_u).$$

At the start of period t, e^{u_t} is known to households, meaning the timing of monetary shocks occurs prior to G&S markets opening.[81] The version of the model with an exogenous money growth rule is referred to as Model v.1.

4.4.5.2. *Inflation targeting monetary policy (Model v.2)*

Forward looking inflation targeting monetary policy takes the form of a Taylor rule specification, as per Taylor (1993), whereby nominal interest rates are adjusted according to the formula:

$$i_t - i^* = \Omega_\pi(E_t(\pi_{t+1}) - \pi^*) + \Omega_y(\ln y_t - \ln y^*) + v_t, \qquad \forall\, t \geq 0 \tag{4.29}$$

$$v_t = \rho_v v_{t-1} + \varepsilon_{v,t}, \qquad |\rho_v| < 1, \qquad \varepsilon_{v,t} \sim N(0, \sigma_v)$$

where Ω_π and Ω_y are inflation and output gap adjustment factors respectively, π^* and y^* are the steady-state values of output and inflation respectively, and v_t is a stochastic monetary policy error. Various calibrations of Ω_π and Ω_y are considered within the discussions that follow.

[81] See footnote 9.

With the inclusion of an inflation targeting rule, the central bank mechanism for achieving the target is to engage in open market operations and thereby control access to liquidity. Therefore, while g_t is set constant and equal to zero, b_t is variable, and the government budget constraint is:

$$\tau_t = m_t - \frac{m_{t-1}}{1+\pi_t} + b_t - (1+i_{t-1})\frac{b_{t-1}}{1+\pi_t}, \qquad \forall\, t \geq 0. \tag{4.30}$$

From Equations (4.24) and (4.30), a positive τ_t implicitly indicates that the treasury issues government bonds, which are purchased by the central bank in exchange for cash. This cash is transferred to firms (in the form of interest bearing loans) through asset managers. Conversely, government borrowing is repaid through negative transfers (taxes), which thereby decrease the money supply. The version of the model with inflation targeting monetary policy is referred to as Model v.2.

4.4.6. Market clearing, general equilibrium, and the steady state

Equilibrium can be defined as a sequence for the endogenous variables satisfying the FOCs derived from the optimisation problems of the representative agents, the various constraints (including the equations governing the G&S production, credit production, states of technology, and monetary policy processes), and the market clearing conditions at every date. The model does not have an analytical solution (due to the functional forms of the equations chosen), but the solutions (also called the policy functions) can be estimated using the method of log-linear approximation around the steady state.[82]

My focus in this chapter is on how the introduction of monetary intermediation (endogenous money) and inflation targeting monetary policy (endogenous monetary policy) influence the dynamics of a DSGE system. This is illustrated in the following section by comparing the volatility and co-movements of various model specifications to various benchmark results.

The results for different specifications of Model v.1 and Model v.2 are presented. Specifications differ according to the choice of values of various parameters. Differences in these parameter values are summarised in Table 4-2 and highlighted in bold italics. There are two specifications of Model v.1 (specifications i and ii), and five specifications of Model v.2 (specifications i to v).

4.5. Results

I present both the long- and short-run properties of the various model specifications. The long-run empirical features are important for assessing how well the steady-state properties of the theoretical model match the data, while the short-run behaviour is important for policy analysis (Walsh, 2010).

[82] I use Dynare version 4.4.3 (Adjemian, et al., 2011) to find the steady state solution of the model.

4.5.1. *Business cycle facts*

Lucas (1977) and Burns & Mitchell (1946) view the business cycle "facts" as the statistical properties of the co-movements of the deviations from trends of various economic aggregates with that of real output. These co-movements are what Lucas regards as the defining feature of the business cycle. Two seminal works aimed at identifying these business cycle "facts" for the US economy are Kydland & Prescott (1990)—hereafter K&P—and Stock & Watson (1999)—hereafter S&W.[83]

K&P report that the major output components tend to move together over the business cycle. The relative standard deviation of investment, and both consumer and producer durables, is several times larger than that of output. The relative standard deviation of spending on nondurable consumption is smaller by a fraction. Additionally, they find little evidence to support views that the price level follows a pro-cyclical pattern (they describe this view as a "monetary myth"), they find little evidence that the real wage is counter-cyclical (or is uncorrelated) with the cycle, and little evidence that the monetary base leads the cycle.

These findings are confirmed and extended by S&W, who find that nominal interest rates are pro-cyclical, real interest rates are slightly counter-cyclical, and that

[83] Additionally McCandless & Weber (1995) focus on the effects of money on the characteristics of economic aggregates across a large sample of countries. They present two key findings. The first is that the correlation between inflation and the growth rate of the money supply is almost unit value, varying depending on the definition of the money supply used. This high correlation, however, does not imply causality, as other factors may generate inflation, with money growing to accommodate higher prices. The second is that there is no correlation between both inflation or money growth, and the growth rate of real output. This supports Taylor's assertion that there is no long-run trade-off between the rate of inflation and the rate of unemployment (Taylor, 1996, p. 186). Walsh (2010, p. 31) writes "the consensus from the empirical literature on the short-run effects of money is that exogenous monetary policy shocks produce hump-shaped movements in real economic activity ... The exact manner in which policy is measured makes a difference, and using an incorrect measure of monetary policy can significantly affect the empirical estimates one obtains".

there are positive correlations between GDP and both inflation and money stocks. Interest in the cyclical behaviour of these time series lies in the *prima facie* inconsistency of this behaviour with established neoclassical growth theory, and the patterns that these time series display. These patterns raise the question of whether they can be replicated in the corresponding statistics of model economies (K&P). These facts and how they compare to early RBC model output from the literature are presented in Table 4-1.

Table 4-1: Stylised business cycle facts

			Results from the literature *					
			K&P	**S&W**	**C&P**		**C&H**	
			US data	US data	US data	Model data	US data	Model data
Variable			(1)	(2)	(3)	(4)	(5)	(6)
Output	Vol.	(1)	1.71	1.66	1.72	1.35	1.72	1.69
	Corr.	(2)	1	1	1	1	1	1
Consumption	Vol.	(3)	0.84	0.78	0.86	0.33	0.86	0.53
	Corr.	(4)	0.76	0.87	0.77	0.84	0.77	0.68
Fixed investment	Vol.	(5)	5.38	4.97	5.34	5.95	8.24	5.9
	Corr.	(6)	0.9	0.89	0.9	0.99	0.91	0.98
Capital stock	Vol.	(7)	0.62					0.38
	Corr.	(8)	-0.08					0.26
Labour Hours	Vol.	(9)	1.47	1.39	1.59	0.77	1.59	1.35
	Corr.	(10)	0.86	0.81	0.86	0.99	0.86	0.99
Real wage	Vol.	(11)	0.91	0.64	0.76			
	Corr.	(12)	0.35	0.16	0.68			
Labour productivity	Vol.	(13)	0.88	1.05	0.9	0.6		0.42
	Corr.	(14)	0.51	0.24	0.41	0.98		0.86
Inflation	Vol.	(15)		0.96			0.57	1.23
	Corr.	(16)		0.15			0.34	-0.14
Nominal interest rate	Vol.	(17)		1.09				0.58
	Corr.	(18)		0.41				-0.01
Real interest rate	Vol.	(19)		0.71				
	Corr.	(20)		-0.28				
Real money base	Vol.	(21)		1.53				
	Corr.	(22)		0.45				
Nominal money base	Vol.	(23)	0.88	1.12			1.46	
	Corr.	(24)	0.41	0.18			0.33	

Notes:
*K&P – Kydland & Prescott (1990); S&W – Stock & Watson (1999); C&P – Cooley & Prescott (1995); C&H – Cooley & Hansen (1995)

Cooley & Prescott (1995)—hereafter C&P—calibrate a divisible labour RBC model subject to persistent stochastic TFP shocks, and conclude that "much but not all variation is accounted for by technology shocks" (Table 4-1, Columns 3 and 4). This is shown by volatility measures in the model data (Column 4) that are generally lower than those of their US data sample measures (Column 3). Additionally, labour input fluctuates only by about half (model volatility of 0.77 compared to volatility in the US data sample of 1.59) suggesting that "important features" of the labour market are not captured. The model does not include any nominal variables.

Cooley & Hansen (1995)—hereafter C&H—specify a CIA RBC model subject to both persistent stochastic TFP shocks and money growth shocks (Table 4-1, Columns 5 and 6). Introducing money into the RBC framework allows for nominal features to be measured. They specify *indivisible* labour in an attempt to better account for the US labour market features. They also distinguish between cash and credit consumption G&S. They find that the co-movements induced by money shocks are inconsistent with those observed in the US data and do not provide "a strong theoretical argument for money shocks being the key impulses driving aggregate fluctuations". Additionally, they were not successful in "accounting for the features of the business cycle that are purely monetary", and thus they conclude that the model does not incorporate features that can account for the properties of the nominal variables over time. Comparing the model results of C&H with those of C&P (Columns 6 and 4 respectively) confirms the findings of Cooley & Hansen (1989). Cooley & Hansen (1989) find that with the inclusion of stochastic money growth shocks, consumption is more variable relative to income (Row 1 shows volatility of 1.69 compared to 1.35), and consumption co-movements with output are smaller (Row 4 shows a correlation coefficient of 0.68 compared to 0.84). With indivisible labour, simulated labour hours are more volatile and better match the features of the US data (Row 9 shows volatility of 1.35 compared to 0.77).

4.5.2. Simulated results of models from within the literature

I replicate the results of both C&P and C&H as well as those of Walsh (2010)—hereafter Walsh—for comparison with the results of the models specified in this study. These three studies provide the benchmark results. I use two different specifications of Walsh's model—one with money growth shocks (specification i) and one without (specification ii)—also for comparative purposes. The parameter values for each of the models are presented in Table 4-2. The simulated results of the various model specifications are presented in Table 4-3.

The replicated results of both C&H and C&P presented in Table 4-3 closely match the original results reported in Table 4-1. Comparing the results of C&H (Table 4-3, Column 2) with those in the Walsh model with money growth shocks (Column 3) confirms that the results are as expected given the differences in parameter values (C&H specify a lower value of ϕ and, unlike Walsh, specify indivisible labour). The model of C&H is more volatile in output (Row 1), consumption (Row 3), and labour hours (Row 9).

Table 4-2: Model parameters

			Parameter values*										
			C&P	C&H	Walsh		Model v.1		Model v.2				
					i	ii	i	ii	i	ii	iii	iv	v
Parameter symbol	Parameter description		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
α	Capital's share of output production	(1)	0.4	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36	0.36
ψ	Pref. param. for leisure**	(2)	0.64	-	1.34	1.34	1.34	1.34	1.34	1.34	1.34	1.34	0.64
η	IES of labour	(3)	1	-	1	1	1	1	1	1	1	1	1
Θ	Pref. param. for cash good purchases	(4)	-	0.84	-	-	-	-	-	-	-	-	-
ω	Disutility of indivisible labour supply	(5)	-	2.53	-	-	-	-	-	-	-	-	-
δ	Depreciation rate	(6)	0.012	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019	0.019
β	Discount factor	(7)	0.987	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989	0.989
Φ	Coeff. of relative risk aversion	(8)	1	1	2	2	2	2	2	2	2	2	1
ρ_z	Persistence of TFP shock	(9)	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
σ_z	TFP shock std .deviation	(10)	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007
ρ_u	Persistence of money growth shock	(11)	-	0.75	0.75	-	0.75	-	-	-	-	-	-
σ_u	Money growth shock std. deviation	(12)	-	0.009	0.009	-	0.009	-	-	-	-	-	-
ξ	Credit scale factor	(13)	-	-	-	-	1.42	1.42	1.42	1.42	1.42	1.42	1.42
γ	Labour's share of credit production	(14)	-	-	-	-	0.21	0.21	0.21	0.21	0.21	0.21	0.21
ρ_v	Persistence of credit shock	(15)	-	-	-	-	0.95	0.95	0.95	0.95	0.95	0.95	0.95
σ_v	Credit shock std. deviation	(16)	-	-	-	-	0.008	0.008	0.008	0.008	0.008	0.008	0.008
Ω_π	Inflation adjustment factor	(17)	-	-	-	-	-	-	1.5	1.5	0.5	1.5	1.5
Ω_y	Output-gap adjustment factor	(18)	-	-	-	-	-	-	-	-	-	0.1	0.01
ρ_υ	Persistence of int.rate shock	(19)	-	-	-	-	-	-	-	0.77	-	-	0.77
σ_υ	Int. rate shock std. deviation	(20)	-	-	-	-	-	-	-	0.004	-	-	0.001

Notes:

* C&P = Cooley & Prescott (1995); C&H = Cooley & Hansen (1995); Walsh = Walsh (2010)

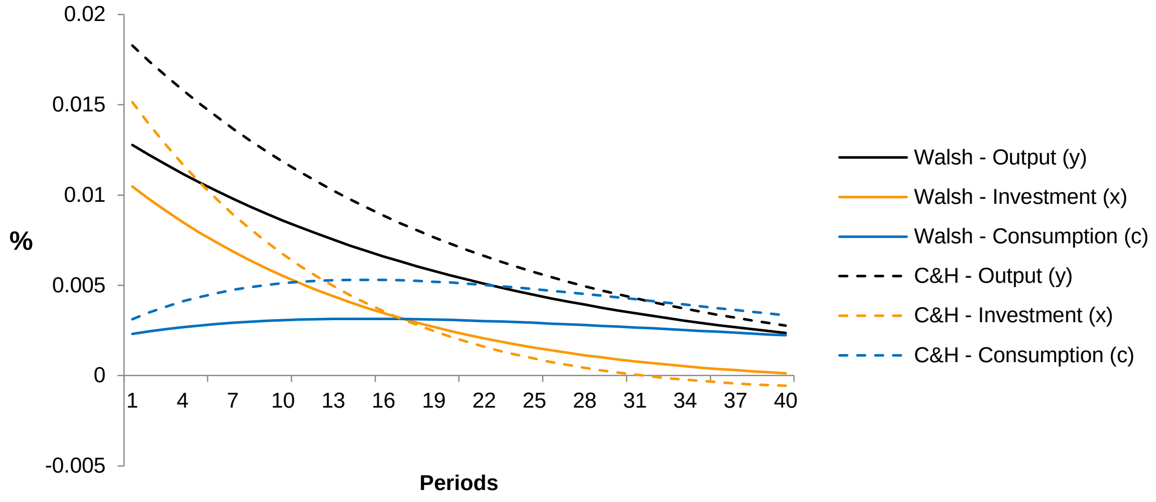
** Where the preference parameter for leisure is below 1, the corresponding preference parameter for consumption is 1-ψ.

Table 4-3: Simulated results of the various model specifications

			Benchmark simulations*				Model simulations						
			C&P	C&H	Walsh		Model v.1		Model v.2				
					i	ii	i	ii	i	ii	iii	iv	v
Variable			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Output	Vol.	(1)	1.37	1.74	1.16	1.16	1.13	1.12	1.11	1.27	1.18	1.25	1.32
	Corr.	(2)	1	1	1	1	1	1	1	1	1	1	1
Consumption	Vol.	(3)	0.32	0.56	0.50	0.29	0.30	0.30	0.29	0.30	0.30	0.29	0.38
	Corr.	(4)	0.89	0.68	0.58	0.94	0.93	0.95	0.95	0.77	0.94	0.93	0.89
Fixed investment	Vol.	(5)	6.02	5.88	4.36	4.20	4.10	4.02	3.97	4.92	4.27	4.67	4.75
	Corr.	(6)	0.99	0.98	0.95	1.00	1.00	1.00	1.00	0.99	1.00	1.00	0.99
Capital stock	Vol.	(7)	0.26	0.39	0.29	0.28	0.27	0.27	0.27	0.31	0.29	0.32	0.32
	Corr.	(8)	0.29	0.31	0.28	0.29	0.29	0.28	0.28	0.29	0.29	0.28	0.32
Labour Hours	Vol.	(9)	0.77	1.38	0.42	0.41	0.45	0.34	0.32	1.04	0.43	0.55	0.68
	Corr.	(10)	0.99	0.99	0.94	0.95	0.82	0.94	0.95	0.72	0.95	0.98	0.95
Real wage	Vol.	(11)	0.61	0.43	0.78	0.78	0.81	0.80	0.81	0.89	0.78	0.73	0.71
	Corr.	(12)	0.99	0.86	0.98	0.99	0.95	0.99	0.99	0.58	0.98	0.99	0.96
Labour productivity	Vol.	(13)	0.61	0.43	0.78	0.78	0.81	0.80	0.81	0.89	0.78	0.73	0.71
	Corr.	(14)	0.99	0.86	0.98	0.99	0.95	0.99	0.99	0.58	0.98	0.99	0.96
Inflation	Vol.	(15)	-	1.28	1.27	0.20	2.53	0.77	0.08	3.78	0.54	3.68	0.75
	Corr.	(16)	-	-0.14	-0.12	-0.50	-0.23	-0.38	0.87	0.22	-0.53	0.38	0.16
Nominal interest rate	Vol.	(17)	-	0.45	0.89	0.02	0.31	0.07	0.09	1.07	0.03	0.14	0.26
	Corr.	(18)	-	-0.02	-0.03	0.97	0.05	0.98	0.97	-0.42	-0.97	-0.98	0.18
Real interest rate	Vol.	(19)	0.03	0.05	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.04
	Corr.	(20)	0.98	0.97	0.97	0.97	0.97	0.97	0.97	0.96	0.97	0.97	0.97
Real money stock	Vol.	(21)	-	0.61	0.50	0.29	2.18	1.02	0.91	6.69	1.62	2.32	1.89
	Corr.	(22)	-	0.62	0.58	0.94	0.59	0.98	0.94	0.60	0.98	0.99	0.73

Notes:
Parameter values are shown in Table 2;
Dynare code is available on request
* C&P = Cooley & Prescott (1995); C&H = Cooley & Hansen (1995); Walsh = Walsh (2010)

Figure 4-1: IRFs to a positive TFP shock in C&H and Walsh



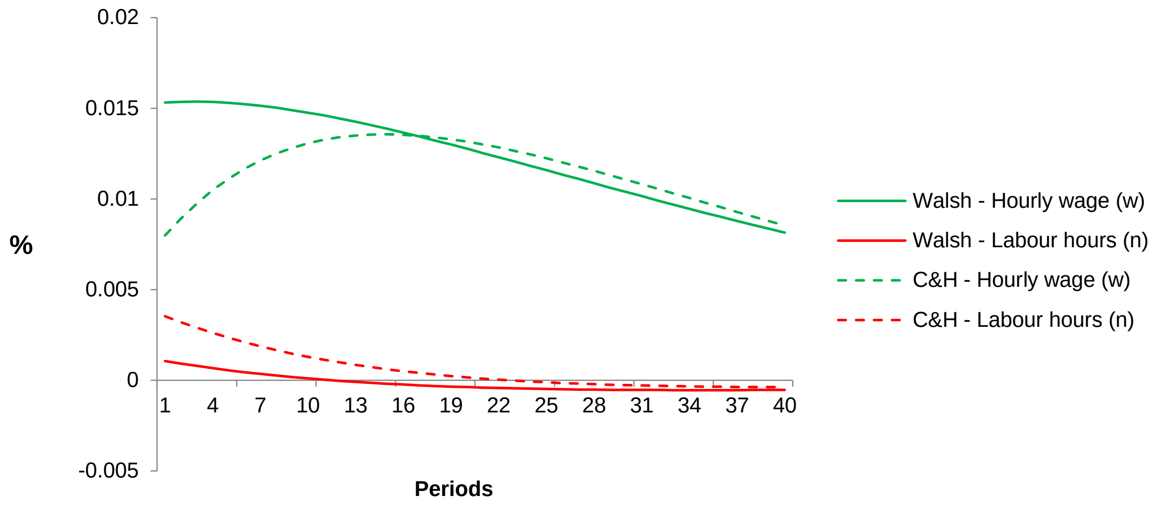
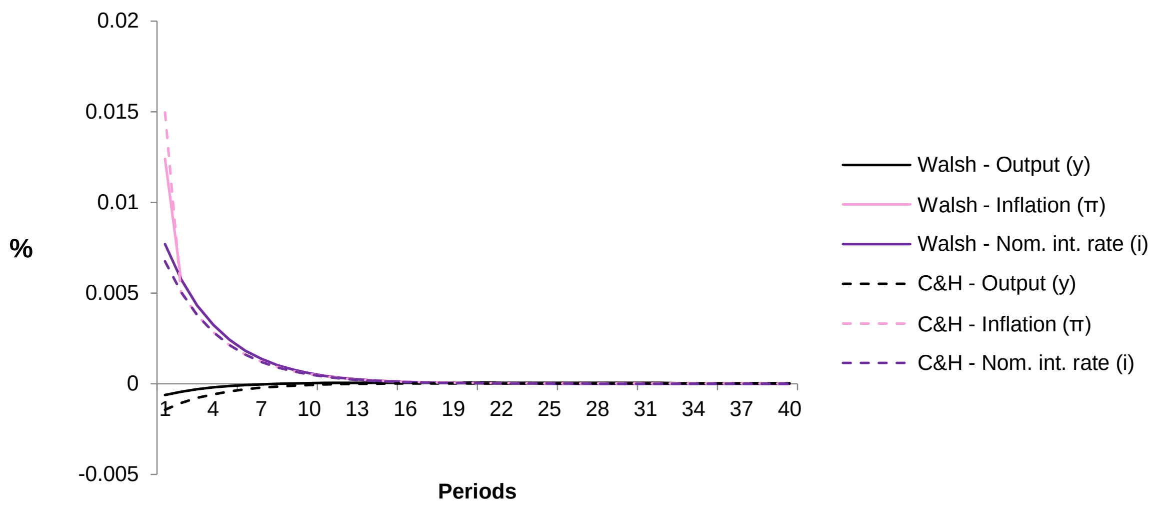


Figure 4-1 illustrates that in response to a positive AR(1) TFP shock, the IRFs of C&H show a higher initial jump in consumption with more hump-shaped dynamics, a higher initial jump in output, investment, and hours worked, and a lower jump in real wages but with a more pronounced hump shape. Money shocks are not neutral in the CIA setup due to the opportunity cost of holding money inherent in the CIA constraint. A popular test of how well a monetary model performs is whether the model can capture the well-known stylised fact known as the "liquidity effect", whereby a positive money supply shock causes a decline in the nominal interest rate along with an increase in output and investment spending (Chen, 2008). Money has no effects on the response of the economy to TFP shocks, but positive money shocks result in higher nominal interest rates and lower output, and the model therefore does not exhibit a liquidity effect (see Figure 4-2). Additionally, the nominal interest rate in the CIA models of both C&H and Walsh is counter-cyclical (Table 4-3, Row 18) which is inconsistent with the US data (see the results of S&W, Table 4-1, Column 2).

Figure 4-2: IRFs to a positive money growth shock in C&H and Walsh



Comparing the features of the Walsh model both with and without money growth shocks (Table 4-3, Columns 3 and 4 respectively) illustrates the impact of including monetary shocks on the business cycle features of the model. The results confirm once again that consumption is more volatile (Row 3) and has smaller co-movements (Row 4). The counter-cyclical nominal interest rate is driven by the inclusion of the exogenous stochastic money growth shocks (Row 18). Inflation exhibits negative co-movement (Row 16) which is also inconsistent with the US data estimates of S&W. Lowering the standard deviation of the money growth shocks removes this counter-cyclical relationship, but substantially underestimates the volatility of these nominal variables (results not presented).

4.5.3. Simulated results of Model v.1 and Model v.2

4.5.3.1. Short-run properties

Figure 4-3 compares the IRFs of Walsh (version ii) to those of both Model v.1 (version ii) and Model v.2 (version i)—that is, models without money growth shocks. Figure 4-3 illustrates that both monetary intermediation and inflation targeting result in smoothing effects compared to the results of Walsh; otherwise the salient nature of the relationships of the various variables remains largely unchanged (other variables not shown).

Figure 4-3: IRFs to a positive TFP shock in Walsh, Model v.1, and Model v.2

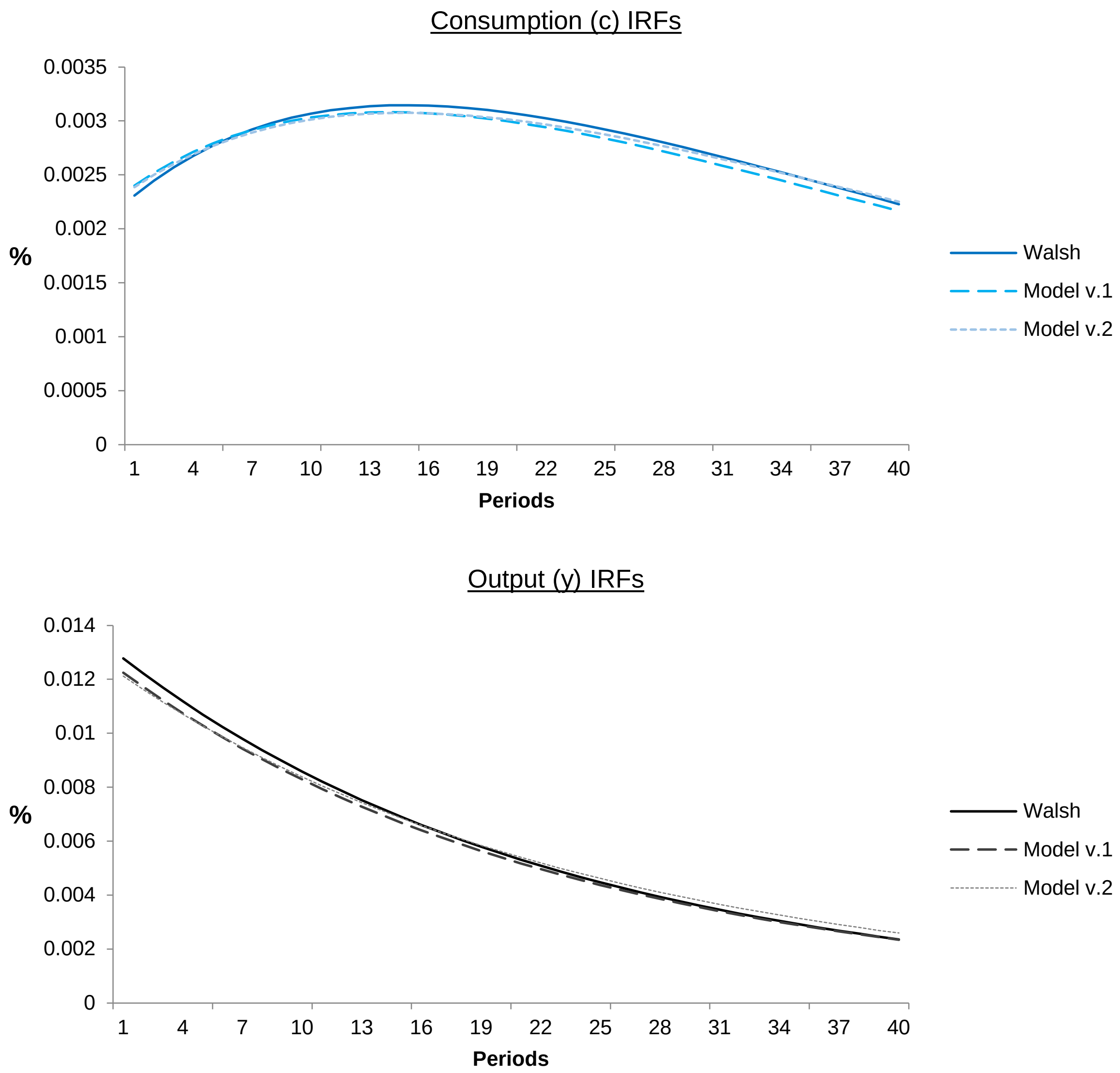
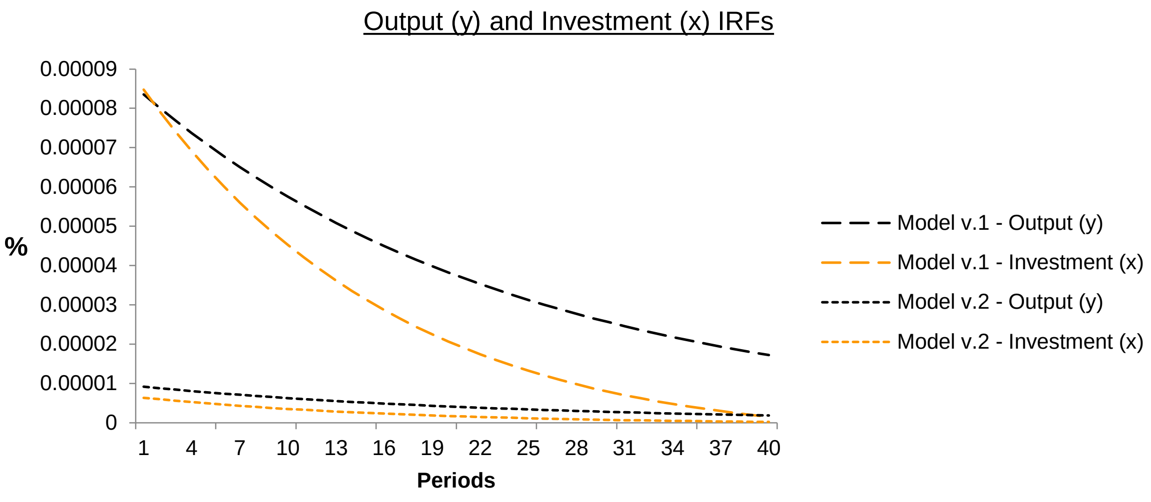
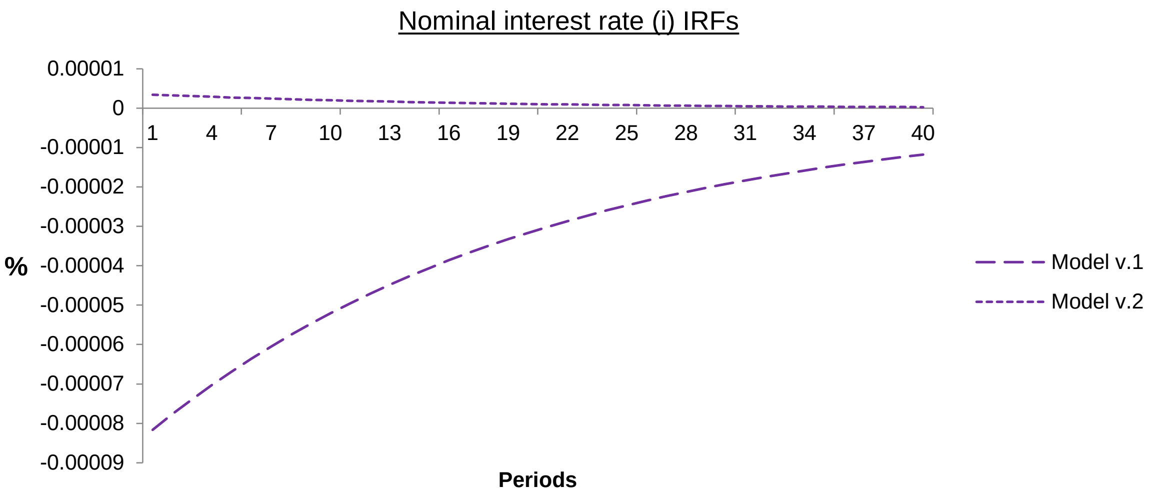


Figure 4-4: IRFs to a positive credit shock in Model v.1 and Model v.2





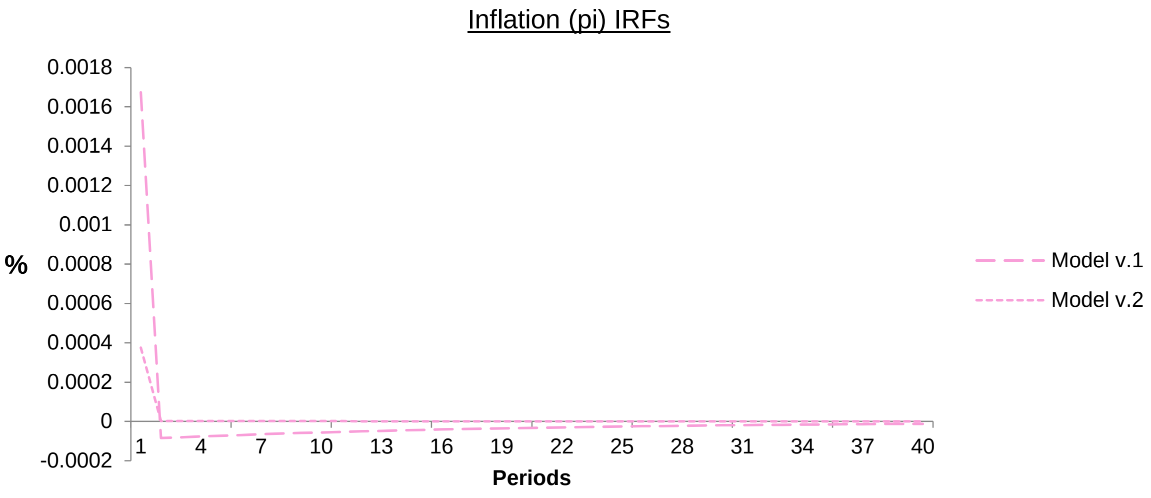


Figure 4-4 compares the effects of credit shocks in Model v.1 (version ii) and Model v.2 (version i). In Model v.1, a positive credit shock increases output and investment, while decreasing the nominal interest rate, and expanding bank deposits and hence endogenous money supply. The model therefore captures the pro-cyclicality associated with financial cycles (Borio, 2014). It is also consistent with a liquidity effect, providing another important improvement in model performance. This provides encouraging support for including bank monetary intermediation as a model component that improves the features of flexible-price DSGE models.

The effects of adding inflation targeting to the model are also shown by comparing the graphs in Figure 4-4. Inflation targeting reduces volatility arising from credit shocks, and leads to a positive relationship between output and nominal interest rates, counteracting the liquidity effect. The results illustrated in Figure 4-4 provide theoretical support for monetary easing as an effective tool for reducing the severity of negative credit shocks. The results suggest that monetary easing has a substantial dampening effect on negative credit shocks, reducing the negative impact on output and minimising the effects of inflation. The model therefore provides theoretical evidence that interest-rate-adjusting inflation targeting can provide a buffer against financial cycle effects.

4.5.3.2. *Long-run properties*

Comparing Table 4-3, Columns 3 and 4 with Columns 5 and 6 indicates that the conclusions of C&H (that the CIA model does not provide a strong argument for exogenous money growth shocks being the key impulses driving aggregate fluctuations) extend to a CIA model augmented with endogenous monetary intermediation. Exogenous money growth shocks contribute little to explaining the observed volatility of output (Row 1) across the benchmark versions of the Walsh model and both the specifications of Model v.1. Relative to the "stylised facts", the co-movements of nominal interest rates (Row 18) are too high in the model versions without money shocks, but the magnitude of these coefficients is too low with the inclusion of exogenous money shocks. By contrast, the credit shocks of Model v.1 greatly improve the volatility of real money stock (Row 21), although the nominal interest rate estimates remain poor relative to the facts.

Adding endogenous inflation targeting monetary policy to the endogenous money model (Model v.2) offers additional encouraging improvements. If money shocks are only in the form of credit shocks (Column 7), both inflation and nominal interest rates are pro-cyclical, although co-movement is overestimated and volatility is underestimated (Table 4-3, Column 7, compared to Table 4-1, Column 2, Rows 15 to 18). These relationships are driven by the policy response to inflation. If money shocks are in the form of both credit shocks and stochastic errors in the setting of interest rates (i.e., $\sigma_{\upsilon} > 0$; Table 4-3, Column 8), then comparing Columns 7 and 8 indicates that: output volatility increases (Row 1); inflation and real money stock volatilities (Rows 15 and 21) increase and are substantially overestimated; and the nominal interest rate (Row

18) is once again counter-cyclical (the inverse relationship between interest rate shocks and output outweighs the positive relationship due to TFP shocks). Additionally, labour hours (Row 9) show increased volatility despite no changes to labour parameters. This has interesting implications suggesting that labour markets are particularly sensitive to errors in monetary policy interest rate setting. Lowering the variance of σ_v reduces these effects.

Various alternative parameterisations of Model v.2 are presented in Columns 9 to 11. Comparing Columns 7 and 9 illustrates that reducing the inflation adjustment factor (Ω_π) increases output volatility (Row 1), and leads to both counter-cyclical nominal interest rates due to the liquidity effect outweighing the policy response (Row 18), and counter-cyclical inflation due to the TFP shock effects outweighing the policy response (Row 16). Alternatively, if the central bank's mandate includes adjustments for output gaps $(\Omega_y > 0)$, comparing Column 7 with Column 10 indicates that economic stability is *reduced*; output volatility increases (Row 1); there is a substantial increase in inflation volatility (Row 15); real money stock volatility increases substantially (Row 21); and the nominal interest rate is counter-cyclical (row 18). Lowering the value of Ω_y tempers these changes (not shown). Finally, the various iterations of the various model specifications indicate that changes to the component parameters estimates can substantially alter the characteristic features of the model economy. Replacing the exogenous money growth rule with endogenous money and an inflation targeting central bank, specifying a low output-gap adjustment parameter, adjusting the model preference parameters to reflect those of C&P and C&H instead of Walsh, and lowering the variance of monetary policy shocks yields the results in Column 11. These results are more in line with literature estimates, and so better capture the salient features of the "facts" reported in Table 4-1, although a number of features remain underestimated. Of particular interest is that these results address empirical shortcomings within RBC theory without reliance on sticky prices or other *ad hoc* frictions.

One glaring shortcoming of the model is that the characteristics of the real interest rate in Table 4-3, Column 11 do not reflect empirical estimates, and offer no improvement over basic CIA estimates. The recipe is clearly missing some key ingredients (perhaps nominal rigidities), and this is an avenue that requires further research. A second avenue for future research includes incorporating indivisible labour to capture extensive margin effects that better match the characteristics of the labour market (and which provide greater amplification of TFP shocks), but this would require modelling a different bank credit technology specification as the current specification requires a divisible labour specification. Habit formation could also be included to improve propagation channels, and capital frictions could be added to make adjustments less smooth. One particularly interesting extension would be to add bank net worth effects along the lines of Carlstrom & Fuerst (1997) in order to introduce richer potential financial cycle effects, and to assess how the hump-shaped dynamics of Carlstrom & Fuerst compare. Wen (2005) provides a discussion on why the propagation mechanisms in RBC models are generally weak, and why the potential modifications mentioned would improve the model results compared to those of empirical estimates.

4.6. Concluding remarks

In this chapter, I examine how both the form of money and the form of the money supply rule affect the nature and amplitude of the business cycle in a flexible-price CIA DSGE model. The model is an extension of the seminal monetary contribution of Cooley & Hansen (1989). I extend the model of Cooley & Hansen to include both endogenous monetary intermediation by banks and endogenous inflation targeting monetary policy rules. I use the extended model to explore the answers to two questions. First, I investigate how monetary intermediation (and credit shocks that change the cost of these monetary intermediation services) impacts on various macroeconomic aggregates. Second, I investigate how these aggregates respond when inflation targeting monetary policy is included in the model.

I find that in response to temporary TFP shocks and relative to benchmark results, both monetary intermediation and inflation targeting monetary policy lower the standard deviations of real output, consumption, and investment. This is due to increased consumption smoothing effects that result from access to monetary intermediation and from inflation targeting interest rate adjustments. In response to credit shocks, inflation targeting is found to lower the standard deviations of real variables resulting from these shocks.

The results provide theoretical support for nominal-interest-rate-adjusting inflation targeting as an effective policy tool for (i) reducing aggregate real economic volatility; and (ii) for reducing the severity of negative credit market shocks which undermine credit provision and output (but with the caveat that if interest rates respond to output gaps, then volatility may in fact increase—this is worth noting given current debate concerning whether central banks should have explicit growth mandates, and the effects of negative interest rates in environments where central banks try to promote both growth and low inflation. These results also mitigate empirical shortcomings within RBC theory without reliance on sticky prices or other *ad hoc* frictions.

Exogenous money growth shocks by themselves contribute little to explaining the observed volatility of output in the model, however, the inclusion of bank intermediated monetary supply greatly amplifies the inflationary effects of these exogenous shocks. This financial accelerator mechanism suggests that liquidity shocks arising from external global financial conditions may be greatly amplified by the domestic banking system (although the model does not provide insight into how exchange rate effects might influence this result). This is an important insight given the current discussion on these matters.

The results suggests that interest-rate-adjusting inflation targeting provides a buffer against financial cycle effects, but because the model is a closed economy, there is no indication of how, for example, carry-trade effects might influence this result. This would be an interesting area of future research. Real interest rate volatility remains largely underestimated in the model, but the volatility of real money stocks is improved. An unexpected outcome of the model is that labour market volatility appears particularly sensitive to errors in monetary policy interest rate setting. It is not clear how this would change if labour were indivisible. Unpacking the effects of this and a number of other parameter changes present useful areas for further research.

Conclusions

> "If old truths are to retain their hold on [people's] minds, they must be restated in the language and concepts of successive generations. What at one time are their most effective expressions gradually become so worn with use that they cease to carry a definite meaning. The underlying ideas may be as valid as ever, but the words, even when they refer to problems that are still with us, no longer convey the same conviction; the arguments do not move in a context familiar to us; and they rarely give us direct answers to the questions we are asking."
>
> - Hayek (2014 [1960])

I. Summary

What do banks do, and what makes them special? Why are they central to the monetary system? I have argued that banks are unique and central to the monetary system because they help people to transact by turning *debt* into *money*. This is different from the idea that banks *lend* money to borrowers *on behalf* of savers, and different to the idea that banks *create money out of nothing*. The differences are important because they have disparate implications for what banks are understood to *do* in the economy, for how banks are understood to *affect* economic systems, and for *why* banks should be regulated. The role of banks in the economy is a renewed topic of interest and debate, and this thesis is a contribution to the debate.

Why do banks turn debt into money? I have argued that banks turn debt into money because people want to be paid with money when selling G&S, but they must often rely on debt to buy G&S. People want to be paid with money because it is less risky and more convenient than dealing with someone else's debt; money is widely accepted, easily transferred, and has easily measured and certain value. Turning debt into money therefore makes it cheaper and easier for people to trade G&S with one another meaning two things. First, banks manage and deal with the risks and inconveniences of the debt buyers agree to in order to buy G&S. Second, they provide sellers who are selling G&S with payment in the form of bank deposit money. This makes banks central to the monetary system.

Why is this theory important for macroeconomics? I have argued that, because banks turn debt into money, the quantity and quality of the debts held on bank balance sheets (bank assets) determines the quantity and quality of bank deposit money supply (bank liabilities). This is important because bank deposit money (and therefore the underlying debt) makes up over 90% of money supply in advanced economies. The 2007/8 financial crisis highlighted how important the quality of this debt is to the monetary system: concerns about poor quality subprime bank assets put the global monetary and financial system under stress, and this stress continues to have lasting negative impacts on the world economy today. The lesson is straightforward: *bank assets matter!* Likewise, in the context of discussing problems in the financial system, Borio (2003) writes: "much of the action is on the *asset side* of [bank] balance sheets". The discussion in this thesis provides a *rich theoretical framework* for understanding why this is the case.

The discussion presented in this thesis indicates that, to the extent market forces are unable to align bank incentives with ensuring efficient bank asset quantity and quality (i.e., to the point where the marginal social cost of transforming a debt into money is greater than the marginal social benefit), regulatory policies that improve the efficiency of bank asset quality and quantity can improve overall economic efficiency. The discussion therefore highlights an important place for prudential regulation.

In summary, I argue that banks are monetary intermediaries. As monetary intermediaries, banks exist to facilitate transactions between buyers and sellers. They improve economic efficiency by providing transaction services that mitigate transaction costs through asset quality and maturity transformation. These services facilitate an efficient intra- and inter-temporal allocation of resources and are therefore critical to the study of *dynamic* macroeconomic systems. The salient features of a bank's monetary intermediation role can be summarised as follows:

(i) bank balance sheets record the value of transactions intermediated by banks;

(ii) these transactions are conducted between buyers and sellers of G&S;

(iii) banks provide credit services to buyers and administer the risks and costs of buyers' debt obligations;

(iv) banks provide liquidity and transaction services to sellers;

(v) bank intermediation services reduce buyer transaction costs, and mitigate risk exposure and other transaction costs to sellers;

(vi) the bank's monetary intermediation services result in solvency and liquidity risk exposure for banks, both of which banks must manage effectively to remain as a going concern; and

(vii) as a result of the services and guarantees that banks provide, bank deposit liabilities provide a generally accepted medium-of-exchange substitute to cash.

The monetary intermediation role of banks is highlighted in early bank literature, but over the past half century, emphasis on this role has receded from the forefront of contemporary bank theory. In this thesis, I highlight the continued relevance of this early literature, and I revive, refine, and expand on the monetary role of banks described in this literature.

Recognising the role of banks as monetary intermediaries has implications for understanding and regulating the role of banks and credit in the monetary system. The first implication is that bank deposits are not created out of nothing. Rather they arise as a result of value-for-value transactions between economic agents intertemporally exchanging G&S on the basis of debt-credit agreements. The second is that the rationale for the existence of a bank and the use of bank liabilities as money is the reduction of the various risks and costs associated with these intertemporal exchanges, and not the transfer of funds from lenders to borrowers. This is a different framing to the role of banks as financial intermediaries. The third is that the quantity and quality of bank-money is underwritten by a portfolio of private debt obligations. A bank's lending decisions therefore impact on the quantity and quality of the money supply and the ability of economic agents to transact. Bank balance sheet assets are therefore central to both broader economic activity and financial stability. Various market failures undermine the efficacy of monetary intermediation by banks, and so regulations that address these failures have the potential to increase economic efficiency. Fourth, from a monetary intermediation perspective, one should expect the growth of bank deposits to be positively correlated with growth in general economic activity. This provides an explanation for the empirical observation that bank intermediation and GDP have strong positive correlations. And finally, monetary intermediation does not depend on prior savings. Instead, it depends on factors that include the need for economic agents to exchange G&S with one another; the ease and willingness to use bank deposit liabilities as the medium-of-exchange; the administrative expertise and ability of banks to manage and reduce the risks and costs related to debt default and deposit withdrawal; and the banking firms' expectations of profitability and solvency with regards to intermediating debt-credit relationships between economic agents. These factors indicate that bank credit extension is likely to be strongly pro-cyclical and are subject to what Keynes' referred to as "animal spirits". To the extent that the business of banking is subject to these mercurial spirits, regulation that is able to "lean against" these pro-cyclical factors is a potential tool for enhancing the resilience and stability of the financial ecosystem. This is increasingly being recognised as the important role of macro-prudential policy (Borio, 2003; BIS, 2018).

II. Key chapter take-aways

Chapter 1

The first chapter identifies key questions and lacunae within bank theory literature, and highlights how the theory of banks as monetary intermediaries provides insight into addressing these issues.

An enduring puzzle from the perspective of *financial* intermediation theory is why do "runnable" bank deposits provide liquidity services when non-bank financial intermediaries, that are not vulnerable to runs, can provide these same liquidity services? Monetary intermediation theory resolves this puzzle. Monetary intermediation theory accounts for the fact that, unlike non-bank financial intermediaries, bank demand deposits are both *easily transferable* and *guaranteed convertible to cash at par and on demand*. These features enable bank deposits to provide transaction services and to function as a direct cash substitute *without liquidation* of the underlying assets. The guarantee makes banks vulnerable to runs under conditions of financial stress. Non-bank financial intermediaries do not provide this guarantee and therefore do not suffer the same vulnerability. The liabilities of non-bank financial intermediaries that are not vulnerable to runs do not function as a cash substitute and so do not provide the same transaction services as banks. Therefore, the liquidity services provided by non-bank financial intermediaries are disparate to the liquidity services provided by banks. This difference is another factor making banks vulnerable to runs. It is therefore clear from a monetary intermediation perspective that non-bank financial intermediaries that are not vulnerable to runs *do not provide the same liquidity services* as banks.

A key question from the perspective of monetary economics is how does the use of money improve resource allocations? Monetary economics focuses on how money enables agents to mitigate various frictions caused by imperfect connections between agents. These frictions limit the ability of agents to credibly commit to terms of trade. Monetary theory shows how, by mitigating these frictions, money expands the set of feasible exchanges, thereby improving resource allocations. Within this broad scope, the theory of monetary intermediation presented in this thesis focuses specifically on how the provision of bank-money improves resource allocations. I show that by transforming debt into money, monetary intermediation expands the set of feasible exchanges by mitigating frictions that limit the *intertemporal* exchange of G&S. This improves resource allocations.

A central theme of early bank theory is exploring how the institutional infrastructure of the banking system enables privately issued debt to provide *new* purchasing power. Early bank theory emphasises how the intermediation of the banking system expands money supply by transforming claims against privately issued debt into money. While this process is discussed in early literature, it is not formally modelled, nor is it applied to a general equilibrium framework. In this thesis, I extend this earlier bank theory by developing a formal model of the process of monetary intermediation by banks, and by incorporating this process into a general equilibrium framework.

Chapter 2

In the second chapter, I develop the concept of monetary intermediation and provide a detailed descriptive account of the role of banks as monetary intermediaries. This provides a rationale for the existence of banks that focuses on the role of banks in reducing transaction costs associated with the exchange of G&S. This advances the theory of banks by fleshing out an alternative framing of a bank's intermediary role. I highlight that, because of this process, banks are exposed to both the default and solvency risks of buyers, and the liquidity risks associated with the cash demands of sellers. It is the bank's ability and skill to manage these risks and to take over the 20-year administration of the debt-credit agreements that justify its existence. This discussion *illustrates how and why a bank's lending decisions impact on both the quantity and quality of an economy's money supp*ly. This is important for unpacking how banks contribute to financial cycles, crises, and general economic stability. This discussion therefore contributes to various debates within banking theory.

In the second part of the chapter, I present a model of monetary intermediation. The model uses an approach inspired by mechanism design methodology to derive the essential risk-sharing characteristics of bank balance sheets (e.g., high leverage and asset portfolios consisting primarily of private debt) and contracts (e.g., guarantees and demand deposits) from first principles. I show that an institution with the characteristics of a bank is an incentive compatible mechanism for exchanging G&S intertemporally by pooling debt on the asset side of its balance sheet, and by providing financial instruments that function as money in the form of a fungible cash substitute on the liabilities side. This *explains why intermediaries with the characteristics of bank* *balance sheets are integral to the monetary system*. The model illustrates that bank balance sheet structure and contract characteristics are a direct consequence of both the role of banks as monetary intermediaries that transform debt into money, and the broader institutional environment (including imperfections and frictions) in which banks perform this role.

The model provides a framework for exploring the role of banks based on the unique monetary transaction services of banks. Monetary intermediation is shown to be more efficient in an environment where frictions, information asymmetries, and other market failures are *reduced*. This result is useful for explaining a number of empirical puzzles such as why financial market development has not coincided with a decrease in bank intermediation, and why the economies of developed nations with more advanced financial markets exhibit higher degrees of bank intermediation as a proportion of GDP. The framework suggests that financial stability and bank resilience is enhanced by measures designed to promote high-quality bank asset portfolios; through measures aimed at reducing undiversifiable risk-taking by banks; through measures to promote financial market development; through measures to reduce information imperfections; and through measures that increase bank capital and collateral requirements.

Chapter 3

In the third chapter, I highlight why an appreciation for framing the role of banks as monetary intermediaries is both theoretically and practically important. I show how framing banks as monetary intermediaries compares and contrasts to framing banks as money creators *ex nihilo*, and framing of banks as financial intermediaries. I illustrate why it is incorrect to interpret modern banking practise as the creation of money *ex nihilo*, and show that this interpretation creates a misleading caricature of a bank's role in the money creation process.

I explore the views of early bank theorists as reflected in the works of Wicksell and Schumpeter, and I show that these authors appreciated and understood the role of banks as risk- and cost-mitigating monetary intermediaries. This discussion is important because understanding the role of banks in the money creation process underpins discussions of the role of banks in the economy. I show that this has implications for macro-model specifications, influencing both the debate and the resulting public policy that is guided by these models. I argue that policy and regulation should be formulated with cognisance of a bank's monetary intermediary role, and I discuss what regulation and policy that recognises a bank's distinct role as monetary intermediary should consider.

I show that, by comparison to financial intermediation, monetary intermediation has empirically disparate effects on the way bank balance sheets respond to and impact on the broader macro-economy.

Chapter 4

The question central to Chapter 4 is how does monetary intermediation affect macroeconomic activity in a general equilibrium framework? I augment a flexible-price DSGE model with monetary intermediary banks and endogenously controlled interest rates set by an inflation-targeting central bank. Banks relax a intertemporal resource constraint by providing a cash equivalent payment option by transforming debt into deposits. I compare how these features influence the characteristics of economic aggregates relative to benchmark models and find improvements both in the business cycle features and in aggregate real volatility.

The chapter makes a number of contributions. By including both endogenous bank intermediated money and endogenous inflation targeting monetary policy, expansionary credit shocks generate a positive liquidity effect, pro-cyclical inflation, and a positive relationship between output and nominal interest rates. These results address empirical shortcomings within standard benchmark RBC models, and indicate that monetary intermediation matters for real economic activity. I find that both monetary intermediation and inflation targeting monetary policy reduce the volatility of real variables in response to temporary productivity shocks. This is due to consumption smoothing effects resulting from access to monetary intermediation services. I show that inflation targeting interest rate adjustments further enhance economic stability by mitigating the impact of negative credit shocks that undermine the provision of monetary intermediation services by banks. These results contribute to current debates centred on the role of the banking sector and monetary policy in the economy more generally.

III. Some additional insights

The framework presented in this thesis has implications across a wide range of subjects. Topics that have been discussed include the link between monetary intermediation and the transmission of monetary policy; the importance of bank transaction services to bank regulation and the prevention of regulatory arbitrage; why bank deposit funding is constrained by a zero lower bound (while non-deposit funding is not); why banks do not substitute away from relatively expensive deposit funding when policy rates are negative; why bank loans are independent of prior savings; why it is misleading to think of bank's creating money out of nothing; why bank intermediation and GDP have strong positive correlations; why banks carry out both lending and demand-deposit functions under one roof; why bank fragility need not be a design feature of banks; and so on. Insights towards additional topics of contemporary interest are considered in the following section.

Bank capital structure and prudential regulation

The capital structure of banks is of interest given growing emphasis on macro-prudential policy (BIS, 2018). The framework presented highlights how bank equity capital provides an internal insurance buffer against otherwise uninsurable risk related to a bank's monetary intermediation function. The implication is that larger capital provisioning is more desirable the more a bank is exposed to uninsurable risk. This highlights the role for bank capital regulations as both a micro- and macro-prudential tool. Monetary intermediation theory therefore provides a rationale for prudential capital requirements as a tool for promoting bank resilience. On the other hand, a trend towards increasing bank leverage ratios is appropriate—in terms of monetary intermediation theory—if this trend is reflective of developments and advances in insurance markets that enable banks to more effectively mitigate risk exposure through off balance sheet insurance.[1] From a monetary intermediation perspective, effective policy making requires that prudential policy makers effectively navigate this tension.

Narrow banks and central bank digital currency

Following the 2007/8 crisis, narrow banking is once again a topic of practical interest as demonstrated by Switzerland's 2018 narrow banking referendum (Dawnay, 2017; Atkins, 2018). Similarly, the Libra Association's 2019 proposal advocates for what is essentially a global narrow bank, although the Libra Association's "deposit" liabilities would be in the form of a digital cryptocurrency (The Libra Association, 2019). Narrow banking requires that the value of a bank's demand deposits are matched in value by the value of cash and government bond assets held by the bank. The underlying motivation for this constraint is to promote stability in the banking system by minimising the risk exposure of bank deposits (Benes & Kumhof, 2012; Pennacchi, 2012; Merton & Thakor, 2018). A similar motivation has been suggested for central banks issuing central bank digital currency (CBDC).[2] When viewed from a monetary intermediation perspective, however, the intended risk reduction seems unlikely. Consider the following examples.

Under a 100% narrow banking regime, for a firm to pay an employee's salary by drawing on a bank overdraft, and for the employee to receive the salary in the form of a

[1] See for example the discussion in Gorton (2012).
[2] See for example the discussion in Engert & Fung (2017).

bank demand-deposit liability, the bank system would need to sell the firm's overdraft to non-bank financial market investors to obtain the cash or government bonds necessary to match the employee's salary deposit. Narrow banking would therefore precipitate a rapid expansion of securitisation markets as banks are forced to take private debt directly to market. Given the events of the recent financial crises there is no clear *prima facie* reason why this growth in securitisation would promote financial stability or reduce moral hazard problems. On the contrary, the 2007/8 crisis suggests that a rapid growth in securitisation is likely to exacerbate these concerns because bank incentives are fundamentally different when bank loans are originated and taken off balance sheet, instead of being retained on balance sheet (Allen, et al., 2015). The risks would remain within the financial system. There would also be a substantial increase in demand for the government issued bills and central bank issued cash needed to underwrite the demand-deposit requirements of the economy. This would likely crowd out the issuance of private commercial debt and inhibit trade financing. This would have far-reaching implications for financial market funding, money market dynamics, and monetary policy implementation.

Under these circumstances, the nature of a bank's various relationships—i.e., monitoring, ameliorating information asymmetries, etc., as discussed by Gorton & Winton (2003)—would be replaced largely by that of financing originator. In this sense, retail and commercial bank operations would come to resemble investment bank operations—a distinction the Glass-Steagall Act aimed to keep separate (White, 1986). In short, moves towards narrow banking would likely be accompanied by rapid growth in activities resembling shadow banking.[3] It is unclear on the basis of the discussions presented in, for example, Benes & Kumhof (2012) or Pennacchi (2012), how narrow banking would address this change in the nature of a bank's relationships, or promote stability in the financial system.

Similar circumstances prevail if central banks issue CBDCs. In this case, a firm may need to pay staff salaries using an overdraft and staff may wish to hold these salary balances in the form of a CBDC deposit. It is not certain what the central banks' matching assets would be, but there are only variations of two possibilities. One possibility is that central banks provide lending facilities directly to firms; i.e., the central bank completely disintermediates the commercial banking system and lends directly to the non-financial corporates and individuals. The second possibility is that central banks act as financial intermediaries between the commercial banking system and the CBDC deposit holders by investing either directly or indirectly in commercial bank liabilities on behalf of CBDC deposit holders. In both cases, central banks take private sector risk onto the asset side of their balance sheets. Either way, central banks would have to match the CBDC account values with assets underwritten by private debt. It is not clear how this reduces aggregate risk within the system.

By contrast, the framework presented in this thesis suggests that originating *safer private* assets, and reducing bank exposure to residual risks through improved provisioning measures (including regulation aimed at providing insurance and building buffers), offers better options for enhancing financial stability than substituting bank

[3] A similar sentiment is expressed in Diamond & Dybvig (1986; 2000).

loans for public debt.[4] This recommendation echoes that of Sissoko (2015b, p. 2), who writes:

> "[l]ike narrow banking proposals, traditional banking theory indicates that structural reform of the current financial system is needed. Unlike narrow banking proposals … the purpose of structural reform … [should be] to regenerate the banking system's ability to originate safe, privately-issued assets."

Bank financing vs financial market financing

In Chapter 2, buyers can access financial markets to sell their debt to investors in exchange for money, or they can approach the banking system to have their debt transformed into money. Banks and financial markets therefore provide two alternative mechanisms of obtaining "purchasing power" from debt; however, the risk distribution differs across these mechanisms and this has regulatory implications.

Debt sold on financial markets is by-and-large issued with a fixed principle and a fixed coupon. A change in the yield occurs with a change in the market price and this leads to the principle value trading at either a premium or a discount—depending on whether the yield has decreased or increased. Investors are therefore exposed to the risk of a variable principle value, and they face the potential of a loss on the principle value if they are forced to liquidate their position prior to the maturity date. By contrast, bank deposit holders face a variable interest rate but face no loss on the principle value of their deposits when liquidating. As a result, for a given change in interest rate, financial

[4] This is consistent with Gorton & Metrick (2010), who state that: "[h]istory has demonstrated two successful methods for regulating privately created money: strict guidelines on collateral, and government guaranteed insurance".

market investors face greater return volatility (which increases the longer the tenor of the debt) than bank deposit holders (whose holdings are priced at immediate tenor). This implies that investors with a potential risk of early liquidation but with higher risk appetite would prefer to purchase debt through financial markets, while investors with lower risk appetite would prefer to hold bank deposits.

We could therefore expect demand deposits to be held by investors with more risk aversion than financial market bond investors. These differences in risk preferences would lead to natural price-discrimination across these mechanisms and imply self-selection through pricing: lower risk bank deposits would be priced at lower interest rates. These statements are empirically testable. Evidence of differing risk preferences would have important implications for bank regulation. It would emphasise the desirability of ensuring that bank deposits are safe. This would support regulation aimed at separating demand deposits from ancillary risk arising from alternative business operations such as investment banking. This would support regulation aimed at separating retail-banking activities from investment-banking activities (as was the case with the Glass-Steagall Act). It would support arguments for improving the riskless nature of deposits through access to lender-of-last-resort facilities or deposit insurance, it would support calls for deposit insurance premiums that are risk sensitive, it would support calls for regulation aimed at promoting bank resilience, and it would support calls for regulation aimed at governing bank asset quality.

IV. Future research

To conclude, three questions arise from the monetary intermediation framework developed in this thesis that have not been explored. These questions frame the scope for future research. The first is a question of the relative size of bank transaction services compared to their portfolio management services in developed versus emerging market economies. As suggested by the results of Chapter 3, these different economic services have different implications for how monetary policy is transmitted, and for how bank regulation is framed and developed. There is therefore value in exploring this question.

Second, it is not clear why in practice cash is such a small fraction of the medium-of-exchange compared to bank deposits. The answer to this puzzle would have implications for monetary policy; for thinking about the optimal size of central bank balance sheets in the age of inflated central bank balance sheets and unconventional monetary policy; for thinking about the benefits and costs of introducing CBDCs; and for questions regarding the funding of fiscal policy (given debate around issues such as the "Green New Deal").

Finally, the framework developed in this thesis offers no real indication of the benefits or the trade-off costs of the various policy interventions that have been discussed. Extensions to the modelling frameworks to formalise the measurement of these aspects are needed to begin assessing the relative net-benefits. This research would be a useful contribution to current interest in developing macro-prudential policy models.

Bibliography

Abel, A. B., 1985. Dynamic Behavior of Capital Accumulation in a Cash-in-Advance Model. *Journal of Monetary Economics,* 16(1), p. 55–71.

Adams, D. W. & Vogel, R. C., 1986. Rural financial markets in low-income countries: Recent controversies and lessons. *World Development,* 14(4), pp. 477-487.

Adjemian, S. et al., 2011. Dynare: Reference Manual, Version 4. *Dynare Working Papers,* Volume 1.

Aikman, D., Haldane, A. G., Hinterschweiger, M. & Kapadia, S., 2018. Rethinking financial stability. *BoE Staff Working Paper No. 712.*

Akerlof, G. A., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics,* 84(3), pp. 488-500.

Algu, Y. & Creamer, K., 2017. Evaluating South Africa's Open Economy. *South African Journal of Economics,* 85(2), pp. 196-221.

Allen, F., Carletti, E. & Gu, X., 2015. The Roles of Banks in Financial Systems. In: A. N. Berger, P. Molyneux & J. O. S. Wilson, eds. *The Oxford Handbook of Banking.* Oxford: Oxford University Press, pp. 27-46.

Allen, F. & Santomero, A., 1998. The theory of financial intermediation. *Journal of Banking & Finance,* Volume 21, pp. 1461-1485.

Allen, F. & Santomero, A. M., 2001. What do financial intermediaries do?. *Journal of Banking & Finance,* Volume 25, pp. 271-294.

Angell, N., 1930. *The Story of Money.* 1st ed. London: Cassell & Company Ltd.

Arena, R. & Festré, A., 1996. *Joseph A.Schumpeter, Historian of Economics.* New York, NY: Routledge.

Arena, R. & Festré, A., 2005. Macroeconomic analysis and individual economic rationality: Some lessons from Wicksell to Schumpeter and von Mises. In:

G. Fontana & R. Realfonzo, eds. *The Monetary Theory of Production: Tradition and Perspectives.* Hampshire: Palgrave Macmillan, pp. 23-38.

Arnon, A., 2011. *Monetary Theory and Policy from Hume and Smith to Wicksell.* Cambridge: Cambridge University Press.

Arrow, K. J. & Debreu, G., 1954. Existence of an equilibrium for a competitive economy. *Econometrica,* 22(3), pp. 265-290.

Atkins, R., 2018. *Radical reform: Switzerland to vote on banking overhaul.* [Online] Available at: https://www.ft.com/content/13b92d86-5810-11e8-bdb7-f6677d2e1ce8 [Accessed 30 May 2018].

Azariadis, C., Bullard, J. & Smith, B. D., 2001. Private and Public Circulating Liabilities. *Journal of Economic Theory,* 99(1-2), pp. 59-116.

Bagehot, W., 1873. *Lombard Street: A Description of the Money Market.* 3rd ed. London: Henry S. King & Co.

Baltensperger, E., 1989. Credit. In: J. Eatwell, M. Milgate & P. Newman, eds. *Money.* London: Palgrave Macmillan, pp. 97-102.

Barrdear, J. & Kumhof, M., 2016. The macroeconomics of central bank issued digital currencies. *Bank of England Staff Working Paper No. 605.*

Baumol, W. J., 1952. The Transactions Demand for Cash: An Inventory Theoretic Approach. *The Quarterly Journal of Economics,* 66(4), pp. 545-556.

Baumol, W. J. & Blinder, A. S., 2011. *Economics: Principles and Policy.* 12th ed. United States: Cengage Learning.

Bech, M. & Garratt, R., 2017. Central bank cryptocurrencies. *BIS Quarterly Review,* pp. 55-70.

Bellofiore, R., 2009. The monetary circuit of capital in the Anti-Critique. In: R. Bellofiore, ed. *Rosa Luxemburg and the Critique of Political Economy.* London: Routledge.

Bellofiore, R. & Realfonzo, R., 2003. Money as Finance and Money as Universal Equivalent: Re-reading Marxian Monetary Theory. In: L. Rochon & S. Rossi, eds.

Modern Theories of Money: The Nature and Role of Money in Capitalist Economies. Cheltenham: Edward Elgar Publishing Limited.

Benes, J. B., Kumhof, M. & Laxton, D., 2014. Financial Crises in DSGE Models: A Prototype Model. *IMF Working Paper WP/14/57.*

Benes, J. & Kumhof, M., 2012. The Chicago Plan Revisited. *IMF Working Paper WP/12/202.*

Benk, S., Gillman, M. & Kejak, M., 2005. Credit shocks in the financial deregulatory era: Not the usual suspects. *Review of Economic Dynamics,* 8(3), pp. 668-687.

Bernanke, B. S., 1986. Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy,* 25(1), pp. 49-99.

Bernanke, B. S., 2012. *Some Reflections on the Crisis and the Policy Response.* New York, Board of Governors of the Federal Reserve System.

Bernanke, B. S. & Blinder, A. S., 1992. The Federal Funds Rate and the Channels of Monetary Transmission. *The American Economic Review,* 82(4), pp. 901-921.

Bernanke, B. S., Gertler, M. & Gilchrist, S., 1999. The Financial Accelerator in a Quantitative Business Cycle Framework. In: J. B. Taylor & M. Woodford, eds. *Handbook of Macroeconomics.* Amsterdam: Elsevier Science B.V., pp. 1341-93.

Besley, T., 1995. Nonmarket Institutions for Credit and Risk Sharing in Low-Income Countries. *The Journal of Economic Perspectives,* 9(3), pp. 115-127.

Best, M. C., Cloyne, J., Ilzetzki, E. & Kleven, H., 2018. Estimating the Elasticy of Intertemporal Substitution Using Mortgage Notches. *NBER Working Paper No. 24948.*

Bhattacharya, S., Boot, A. W. A. & Thakor, A. V., 1998. The Economics of Bank Regulation. *Journal of Money, Credit and Banking,* 30(4), pp. 745-770.

Bhattacharya, S. & Thakor, A. V., 1993. Contemporary Banking Theory. *Journal of Financial Intermediation,* Volume 3, pp. 2-50.

Bindseil, U., 2004. The Operational Target of Monetary Policy and the Rise and Fall of Reserve Position Doctrine. *ECB Working Paper No. 372.*

BIS, 2018. *BIS Annual Economic Report,* Basel: Bank for International Settlements.

Black, F., 1970. Banking and Interest Rates in a World Without Money: The effects of Uncontrolled Banking. *Journal of Bank Research,* Volume 1, pp. 8-20.

Blanchard, O., Dell'Ariccia, G. & Mauro, P., 2010. Rethinking Macroeconomic Policy. *Journal of Money, Credit and Banking,* Supplement to Vol. 42,(6), pp. 199-215.

Boianovsky, M., 2002. Simonsen and the early history of the cash-in-advance. *The European Journal of the History of Economic Thought,* 9(1), pp. 57-71.

Borio, C., 2003. Towards a macroprudential framework for financial supervision and regulation?. *BIS Working Papers No 128.*

Borio, C., 2014. The financial cycle and macroeconomics: What have we learnt?. *Journal of Banking & Finance,* Volume 45, pp. 182-198.

Borio, C., 2019. On money, debt, trust and central banking. *BIS Working Papers No 763.*

Borio, C. & Disyatat, P., 2011. Global imbalances and the financial crisis: Link or no link?. *BIS Working Papers No 346.*

Boyd, J. H. & Prescott, E. C., 1986. Financial intermediary-coalitions. *Journal of Economic Theory,* 38(2), pp. 211-232.

Brunner, K. & Meltzer, A. H., 1971. The Uses of Money: Money in the Theory of an Exchange Economy. *The American Economic Review,* 61(5), pp. 784-805.

Brunnermeier, M. K. & Sannikov, Y., 2016. The I Theory of Money. *Working Paper,* August.

Bryant, J., 1980. Transaction Demand for Money and Moral Hazard. In: J. Kareken & N. Wallace, eds. *Models of Monetary Economies: Proceedings and Contributions from Patricipants of a December 1978 Conference Sponsored by the Federal Reserve Bank of Minneapolis.* Minneapolis:Federal Reserve Bank of Minneapolis, pp. 233-241.

Bullard, J. & Smith, B. D., 2003a. The value of inside and outside money. *Journal of Monetary Economics,* pp. 389-417.

Bullard, J. & Smith, B. D., 2003b. Intermediaries and payments instruments. *Journal of Economic Theory,* 109(2), pp. 172-197.

Burns, A. F. & Mitchell, W. C., 1946. *Measuring Business Cycles.* New York: NBER Book Series Studies in Business Cycles.

Calomiris, C. W. & Kahn, C. M., 1991. The Role of Demandable Debt in Structuring Optimal Banking Arrangements. *The American Economic Review,* 81(3), pp. 497-513.

Campbell, J. Y., 1994. Inspecting the mechanism: An analytical approach to the stochastic growth model. *Journal of Monetary Economics,* Volume 33, pp. 463-506.

Campbell, J. Y. & Mankiw, G. N., 1989. Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence. *NBER Macroeconomics Annual,* Volume 4, pp. 185-216.

Caprio, G. J. & Honohan, P., 2015. Banking Crises: Those Hardy Perennials. In: A. N. Berger, P. Molyneux & J. O. S. Wilson, eds. *The Oxford Handbook of Banking.* Oxford: Oxford University Press, pp. 700-720.

Carlstrom, C. T. & Fuerst, T. S., 1997. Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis. *The American Economic Review,* 87(5), pp. 893-910.

Cencini, A., 2002. *Monetary Macroeconomics: A New Approach.* London: Routledge.

Cerf, B., 1944. *Try and Stop Me.* New York: Simon & Schuster.

Cesarano, F., 2014. The Puzzle of Metallism: Searching for the Nature of Money. *History of Political Economy,* 46(2), pp. 177-210.

Chen, S.-S., 2008. The Liquidity Effect in a Flexible-Price Monetary Model. *Oxford Economic Papers,* 60(1), pp. 122-142.

Christensen, I. & Dib, A., 2008. The financial accelerator in an estimated New Keynesian model. *Review of Economic Dynamics,* Volume 11, pp. 155-178.

Christiano, L. J. & Eichenbaum, M., 1992. Current Real Business Cycle Theories and Aggregate Labor-Market Fluctuations. *American Economic Review,* 82(3), pp. 430-450.

Clower, R. W., 1967. A Reconsideration of the Microfoundations of Monetary Theory. *Western Economic Journal,* 6(1), pp. 1-8.

Clower, R. W., 1970. Is There an Optimal Money Supply?. *The Journal of Finance,* 25(2), pp. 425-433.

Coase, R., 1937. The Nature of the Firm. *Economica,* pp. 386-405.

Cooley, T. F. & Hansen, G. D., 1989. The Inflation Tax in a Real Business Cycle Model. *The American Economic Review,* 79(4), pp. 733-748.

Cooley, T. F. & Hansen, G. D., 1995. Money and the Business Cycle. In: T. F. Cooley, ed. *Frontiers of Business Cycle Research.* Princeton, New Jersey: Princeton University Press, pp. 175-216.

Cooley, T. F. & LeRoy, S. F., 1985. Atheoretical Macroeconometrics: A Critique. *Journal of Monetary Economics,* Volume 16, pp. 283-208.

Cooley, T. F. & Prescott, E. C., 1995. Economic Growth and Business Cycles. In: T. F. Cooley, ed. *Frontiers of Business Cycle Research.* Princeton, New Jersey: Princeton University Press, pp. 1-38.

Coppola, F., 2012. *I am a bank.* [Online]
Available at: http://coppolacomment.blogspot.com/2012/07/i-am-bank.html

Corbae, D. & Levine, R., 2018. Competition, Stability, and Efficiency in Financial Markets. *Kansas City Fed Working Papers.*

Coval, J. D. & Thakor, A. V., 2005. Financial intermediation as a beliefs-bridge between optimists and pessimists. *Journal of Financial Economics,* Volume 75, pp. 535-569.

Daastøl, A. M., 2015. Austerity Versus Productive Investment: Two Traditions in Capital Formation and Growth. In: J. G. Backhaus, ed. *Great Nations at Peril.* Switzerland: Springer International Publishing.

Dang, T. V., Gorton, G. & Holmström, B., 2015. Ignorance, Debt and Financial Crises. *Draft working paper.*

Dang, T. V., Gorton, G., Holmström, B. & Ordoñez, G., 2017. Banks as Secret Keepers. *American Economic Review,* 107(4), p. 1005–1029.

Danthine, J.-P., 1997. In Search of a Successor to IS-LM. *Oxford Review of Economic Policy,* 13(3), pp. 135-144.

Davies, C., Gillman, M. & Kejak, M., 2012. Deriving the Taylor Principle when the Central Bank Supplies Money. *Cardiff Economics Working Papers No. E2012/20.*

Davies, G., 2010. *A History of Money.* 3rd ed. Cardiff: University of Wales Press.

Dawnay, E., 2017. *The Background to the National Referendum on Sovereign Money in Switzerland,* Wettingen: Monetäre Modernisierung.

de Boyer, J., 1985. Taux d'intérêt et quantité de monnaie: note sur la distinction entre deux sphères de circulation chez T. Tooke, K. Wicksell, J.M. Keynes et Schumpeter. In: *Production, circulation et monnaie.* Paris: Presse Universitaires de France, pp. 371-386.

de Roover, R., 1948. *Money, Banking and Credit in Mediaeval Bruges.* Cambridge Massachusetts: The Mediaeval Academy of America.

de Soto, J. H., 2009. *Money, Bank Credit, and Economic Cycles.* 2nd ed. Auburn, Alabama: Ludwig von Mises Institute.

DeAngelo, H. & Stulz, R. M., 2015. Liquid-claim production, risk management, and bank capital structure: Why high leverage is optimal for banks. *Journal of Financial Economics,* 116(2), pp. 219-236.

Decker, F. & Goodhart, C. A., 2018. Credit mechanics – a precursor to the current money supply debate. *CEPR discussion paper series DP13233.*

Diamond, D. & Dybvig, P., 2000. Bank Runs, Deposit Insurance, and Liquidity. *Federal Reserve Bank of Minneapolis Quarterly Review,* 24(1), pp. 14-23.

Diamond, D. W. & Dybvig, D. H., 1986. Banking Theory, Deposit Insurance, and Bank Regulation. *The Journal of Business,* 59(1), pp. 55-68.

Diamond, D. W. & Dybvig, P. H., 1983. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy,* 91(3), pp. 401-419.

Diamond, P., 1984. Money in Search Equilibrium. *Econometrica,* 52(1), pp. 1-20.

Dittmar, R. D., Gavin, W. T. & Kydland, F. E., 2005. Inflation Persistence and Flexible Prices. *International Economic Review,* 46(1), pp. 245-261.

Donaldson, J. R., Piacentino, G. & Thakor, A., 2016. Warehouse Banking. *Working Paper,* April .

Drechsler, I., Savov, A. & Schnabl, P., 2017. The Deposits Channel of Monetary Policy. *The Quarterly Journal of Economics,* 132(4), pp. 1819-76.

Drehmann, M., Claudio Borio, C. B. E. V. & Tsatsaronis, K., 2012. Characterising the Financial Cycle: Don't Lose Sight of the Medium Term!. *BIS Working Paper No. 380.*

Duistermaat, K., 2010. Administration in Neolithic Societies? The first use of seals in Syria and some considerations on seal owners, seal use and private property. In: W. Müller, ed. *Die Bedeutung der minoischen und mykenischen Glyptik. VI. Internationales Siegel-Symposium, Marburg, 9.–12. Oktober 2008.* Mainz am Rhein: Philipp von Zabern, pp. 167-182.

Eggertsson, G. B., Juelsrud, R. E., Summers, L. H. & Wold, E. G., 2019. Negative Nominal Interest Rates and the Bank Lending Channel. *NBER Working Paper No. 25416.*

Ekelund, Jr., R. B., Hébert , R. F. & Tollison, R. D., 1989. An Economic Model of the Medieval Church: Usury as a Form of Rent Seeking. *Journal of Law, Economics, & Organization,* 5(2), pp. 307-331.

Elmendorf, D. W., 1996. The effect of interest-rate changes on household saving and consumption: a survey. *Finance and Economics Discussion Series 96-27, Board of Governors of the Federal Reserve System (US).*

Engert, W. & Fung, B. S.-C., 2017. Central bank digital currency: Motivations and implications. *Bank of Canada Staff Discussion Paper, No. 2017-16.*

Engert, W. & Fung, B. S. C., 2017. Central Bank Digital Currency: Motivations and Implications. *Bank of Canada Staff Discussion Paper 2017-16.*

Fama, E., 1980. Banking in the Theory of Finance. *Journal of Monetary Economics,* Volume 6, pp. 39-57.

Ferguson, N., 2009. *The Ascent of Money: A Financial History of the World.* London: Penguin Books Ltd.

Festré, A., 2002. Money, Banking and Dynamics: Two Wicksellian routes from Mises to Hayek and Schumpeter. *The American Journal of Economics and Sociology,* 61(2), pp. 439-480.

Festré, A. & Nasica, E., 2009. Schumpeter on money, banking and finance: An institutionalist perspective. *The European Journal of the History of Economic Thought,* 19(2), pp. 325-356.

Fisher, I., 1933. The Debt-Deflation Theory of Great Depressions. *Econometrica,* 1(4), pp. 337-357.

Flanders, M. J., 1990. A treatise on money. In: *International monetary economics, 1870–1960: Between the Classical and the New Classical.* New York: Cambridge University Press, pp. 170-185.

Flannery, M., 1994. Debt Maturity and the Deadweight Cost of Leverage: Optimally Financing Banking Firms. *The American Economic Review,* 84(1), pp. 320-331.

Fontana, G., 2006. The "New Consensus" View of Monetary Policy: A New Wicksellian Connection?. *The Levy Economics Institute Working paper 476.*

Fontana, G., 2007. Why money matters: Wicksell, Keynes, and the new consensus view on monetary policy. *Journal of Post Keynesian Economics,* 30(1), pp. 43-60.

Frangipane, M., 2016. The Origins of administrative practices and their developments in Greater Mesopotamia. The evidence from Arslantepe. *Archéo-Nil,* Volume 26, pp. 9-32.

Freixas, X. & Rochet, J.-C., 2008. *Microeconomics of Banking.* 2nd ed. Cambridge, Massachusetts: MIT Press.

Frisch, R., 1967 [1933]. Propagation problems and impulse problems in dynamic economics. In: *Economic Essays in Honour of Gustav Cassel.* s.l.:Frank Cass & Co. Ltd, pp. 171-205.

Galati, G. & Moessner, R., 2013. Macroprudential Policy - A literature review. *Journal of Economic Surveys,* 27(5), pp. 846-878.

Gardner, P., 1918. *A History of Ancient Coinage 700-300 B.C.*. Oxford: Clarendon Press.

Gillet Bros. Discount Co. Ltd, 1964. *The Bill on London: or the Finance of Trade by Bills of Exchange.* London: Chapman & Hall.

Gillman, M. & Kejak, M., 2011. Inflation, Investment and Growth: a Money and Banking Approach. *Economica,* 78(310), p. 260–282.

Gillman, M., 1993. The welfare cost of inflation in a cash-in-advance economy with costly credit. *Journal of Monetary Economics,* Volume 31, pp. 97-115.

Gillman, M. & Kejak, M., 2005. Inflation and Balanced-Path Growth with Alternative Payment Mechanisms. *The Economic Journal,* Volume 115, pp. 247-270.

Gillman, M., Siklos, P. L. & Silver, J. L., 1997. Money Velocity with Costly Credit. *Journal of Economic Research,* Volume 2, pp. 179-207.

Giovannini, A., 1983. The interest elasticity of savings in developing countries: The existing evidence. *World Development,* 11(7), pp. 601-607.

Gnos, C., 2006. French circuit theory. In: P. Arestis & M. Sawyer, eds. *A Handbook of Alternative Monetary Economics.* Cheltenham: Edward Elgar Publishing Limited.

Goodhart, C. A., 1998. The two concepts of money: Implications for the analysis of optimal currency areas. *European Journal of Political Economy,* Volume 14, pp. 407-432.

Goodhart, C. A. E., 2009. The Continuing Muddles of Monetary Theory: A Steadfast Refusal to Face Facts. *Economica,* Volume 76, pp. 821-30.

Goodhart, C. & Jensen, M., 2015. Currency School versus Banking School: An ongoing confrontation. *Economic Thought,* 4(2), pp. 20-31.

Goodspeed, T. B., 2012. The Wicksell Connection. In: *Rethinking the Keynesian Revolution: Keynes, Hayek, and the Wicksell Connection.* New York: Oxford University Press.

Gorton, G., 2012. *Misunderstanding Financial Crises: Why We Don't See Them Coming.* Oxford, England: Oxford University Press.

Gorton, G., 2014. Some Reflections on the Recent Financial Crisis. In: R. Acharyya & S. Marjit, eds. *Trade, Globalization and Development: Essays in Honor of Kalyan K. Sanyal.* New Delhi: Springer, pp. 161-184.

Gorton, G. & Metrick, A., 2010. Regulating the Shadow Banking System. *Brookings Papers on Economic Activity,* pp. 261-312.

Gorton, G. & Pennacchi, G., 1990. Financial Intermediaries and Liquidity Creation. *The Journal of Finance,* 45(1), pp. 49-71.

Gorton, G. & Winton, A., 2003. Financial Intermediation. In: G. M. Constantinides, M. Harris & R. Stulz, eds. *Handbook of the Economics of Finance.* Amsterdam: North-Holland, pp. 431-552.

Gottschalk, J., 2001. An Introduction into the SVAR Methodology: Identification, Interpretation and Limitations of SVAR models. *Kiel Working Paper No. 1072.*

Graeber, D., 2011. *Debt: The first 5000 years.* New York: Melville House.

Grandmont, J. & Younes, Y., 1972. On the Role of Money and the Existence of a Monetary Equilibrium. *Review of Economic Studies,* Volume 39, p. 355–372.

Graziani, A., 1989. The Theory of the Monetary Circuit. *Thames Papers in Political Economy,* Issue Spring, pp. 1-26.

Graziani, A., 2003. *The Monetary Theory of Production.* Cambridge: Cambridge University Press.

Greenbaum, S. I., Thakor, A. V. & Boot, A. W. A., 2016. *Contemporary Financial Intermediation.* 3rd ed. Amsterdam: Elsevier Academic Press.

Greif, A., 1993. Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review,* 83(3), pp. 525-548.

Gu, C., Mattesini, F. & Wright, R., 2016. Money and credit redux. *Econometrica,* 84(1), pp. 1-32.

Gu, C., Monnet, C., Nosal, E. & Wright, R., 2019. On the Instability of Banking and Financial Intermediation. *Working paper,* April.

Gurley, J. G. & Shaw, E. S., 1960. *Money in a Theory of Finance.* s.l.:Brookings Institution.

Hagemann, H., 2015. L. Albert Hahn's Economic Theory of Bank Credit. *Journal of Post Keynesian Economics,* 37(2), pp. 309-335.

Hahn, F. H., 1973. On the Foundations of Monetary Theory. In: M. Parkin & A. R. Nobay, eds. *Essays in Modern Economics.* s.l.:Longman, pp. 230-242.

Hahn, L. A., 2015 [1920]. *Economic Theory of Bank Credit.* Oxford: OUP.

Hall, R. E., 1988. Intertemporal Substitution in Consumption. *Journal of Political Economy,* 96(2), pp. 339-357.

Hansen, A. H., 1949. *Monetary Theory and Fiscal Policy.* Tokyo: McGraw-Hill Kogakusha.

Hansen, L. P., 2014. Challenges in Identifying and Measuring Systemic Risk. In: M. Brunnermeier & A. Krishnamurthy, eds. *Risk Topography: Systemic Risk and Macro Modeling.* s.l.:University of Chicago Press, pp. 15-30.

Hassan, S., 2015. Speculative Flows, Exchange Rate Volatility and Monetary Policy: The South African Experience. *South African Reserve Bank Working Paper Series WP/15/02.*

Haubrich, J. G. & King, R. G., 1990. Banking and insurance. *Journal of Monetary Economics,* 26(3), pp. 361-386.

Hawtrey, R. G., 1919. *Currency and Credit.* London: Longmans, Green and Co..

Hayek, F. A., 2014 [1960]. *The Constitution of Liberty.* London: Routledge.

Heinsohn, G. & Steiger, O., 2000. The property theory of interest and money. In: J. Smithin, ed. *What is Money?.* London: Routledge.

He, P., Huang, L. & Wright, R., 2005. Money and Banking in Search Equilibrium. *International Economic Review,* pp. 637-670.

Hicks, J. R., 1935. A Suggestion for Simplifying the Theory of Money. *Economica,* 2(5), pp. 1-19.

Holmström, B., 2015. Understanding the role of debt in the financial system. *BIS Working Papers No 479.*

Holmström, B. & Tirole, J., 1997. Financial Intermediation, Loanable Funds, and the Real Sector. *The Quarterly Journal of Economics,* 112(3), pp. 663-691.

Holtfrerich, C.-L., 1989. Zur Entwicklung der monetären Konjunkturtheorien: Wicksell, Schumpeter, Hahn, Mises und Hayek [On the Development of Monetary Busines Cycle Theories: Wicksell, Schumpeter, Hahn, Mises and Hayek]. In: B. Schefold, ed. *Studien zur Entwicklung der ökonomischen Theorie.* Berlin: Duncker & Humblot, pp. 103-140.

Hudson, M., 2004. The Archaeology of Money: Debt versus Barter Theories of Money's Origins. In: L. R. Wray, ed. *Credit and State Theories of Money: The Contributions of A. Mitchell Innes.* Cheltenham: Edward Elgar Publishing, pp. 99-127.

Hurwicz, L., 1973. The Design of Mechanisms for Resource Allocation. *The American Economic Review,* 63(2), pp. 1-30.

Hurwicz, L. & Reiter, S., 2006. *Designing Economic Mechanisms.* Cambridge University Press: Cambridge.

Inada, K.-I., 1963. On a Two-Sector Model of Economic Growth: Comments and a Generalization. *The Review of Economic Studies,* 30(2), pp. 119-127.

Ingham, G., 2000. 'Babylonian madness': On the historical and sociological origins of money. In: J. Smithin, ed. *What is Money?.* London: Routledge, pp. 16-41.

Innes, A. M., 1913. What is Money. *Banking Law Journal,* 30(5), pp. 377-408.

Innes, A. M., 1914. The Credit Theory of Money. *The Banking Law Journal,* Volume 31, pp. 151-168.

Ireland, P. N., 2008. The Monetary Transmission Mechanism. In: L. Blume & S. Durlauf, eds. *The New Palgrave Dictionary of Economics.* 2nd ed. Hampshire: Palgrave Macmillan Ltd.

Jacklin, C. J., 1987. Demand Deposits, Trading Restrictions, and Risk Sharing. In: E. Prescott & N. Wallace, eds. *Contractual Arrangements for Intertemporal Trade.* s.l.:University of Minnesota Press, pp. 26-47.

Jakab, Z. & Kumhof, M., 2015. Banks are not intermediaries of loanable funds — and why this matters. *Bank of England Working Paper No. 529.*

Julin, E., 2017. *The Riksbank's e-krona project,* Stockholm: Sveriges Riksbank.

Kagan, D., 1982. The Dates of the Earliest Coins. *American Journal of Archaeology,* 86(3), pp. 343-360.

Kahn, C. M. & Roberds, W., 2007. Transferability, finality, and debt settlement. *Journal of Monetary Economics,* 54(4), pp. 955-978.

Kahn, C. M. & Roberds, W., 2009. Why pay? An introduction to payments economics. *Journal of Financial Intermediation,* 18(1), pp. 1-23.

Kaldor, N., 1970. The New Monetarism. *Lloyds Bank Review,* pp. 1-18.

Karagedikli, Ö. & Matheson, T., 2010. RBCs and DSGEs: The Computational Approach to Business Cycle Theory and Evidence. *Journal of Economic Surveys,* 24(1), p. 113–136.

Kashyap, A. K., Rajan, R. & Stein, J. C., 2002. Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking. *The Journal of Finance,* 57(1), pp. 33-73.

Kashyap, A. K. & Stein, J. C., 1995. The impact of monetary policy on bank balance sheets. *Carnegie-Rochester Conference Series on Public Policy,* Volume 42, pp. 151-195.

Kashyap, A. K. & Stein, J. C., 2000. What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?. *The American Economic Review,* 90(3), pp. 407-428.

Kehoe, T. & Levine, D., 1993. Debt-Constrained Asset Markets. *Review of Economic Studies,* 60(4), pp. 865-888.

Keynes, J. M., 1930. *A Treatise on Money Volumes I & II.* London: Macmillan.

Kilian, L., 2013. Structural vector autoregressions. In: N. Hashimzade & M. A. Thornton, eds. *Handbook of Research Methods and Applications in Empirical Macroeconomics.* Cheltenham, UK: Edward Elgar, pp. 515-554.

Kindleberger, C. P., 1984. *A Financial History of Western Europe.* London: George Allen & Unwin (Publishers) Ltd.

King, R. G., Plosser, C. I. & Rebelo, S. T., 1988. Production, Growth and Business Cycles I: The Basic Neoclassical Model. *Journal of Monetary Economics,* Volume 21, pp. 195-232.

King, R. & Levine, R., 1993. Finance and growth: Schumpeter might be right. *Quarterly Journal of,* Volume 108, p. 713–737.

Kiyotaki, N. & Moore, J., 2002. Evil is the Root of All Money. *American Economic Review,* 92(2), pp. 62-66.

Kiyotaki, N. & Wright, R., 1989. On Money as a Medium of Exchange. *Journal of Political Economy,* 97(4), pp. 927-954.

Knapp, G. F., 1924. *The State Theory of Money.* London: Macmillan & Co. Ltd..

Knight, F. H., 1964 [1921]. *Risk, Uncertainty and Profit.* New York: Reprints of Economic Classics.

Kocherlakota, N., 1998. Money is Memory. *Journal of Economic Theory,* October, Volume 81, pp. 232-251.

Kohn, M., 1984. The Finance (Cash-in-Advance) Constraint Comes of Age: A Survey of Some Recent Developments in the Theory of Money. *Dartmouth College Working Paper No. 84-1.*

Kregel, J. A., 1986. Shylock and Hamlet or Are there Bulls and Bears in the Circuit?. *Economies et Sociétés,* Volume 20, pp. 11-22.

Kurke, L., 1999. *Coins, Bodies, Games, and Gold.* Princeton NJ: Princeton University Press.

Kydland, F. E. & Prescott, E. C., 1990. Business Cycles: Real Facts and a Monetary Myth. *Federal Reserve Bank of Minneapolis Quarterly Review,* 14(2), pp. 3-18.

Lagos, R., Rocheteau, G. & Wright, R., 2017. Liquidity: A New Monetarist Perspective. *Journal of Economic Literature,* 55(2), pp. 371-440.

Laidler, D., 2009. Financial Stability, Monetarism and the Wicksell Connection. *Review of Economic Analysis,* Volume 1, pp. 60-79.

Leijonhufvud, A., 1979. The Wicksell Connection: Variations on a Theme. *UCLA Department of Eocnomics Working paper No.165.*

Levine, R. & Zervos, S., 1998. Stock markets, banks, and economic growth. *American Economic Review,* Volume 88, p. 537–558.

Lewis, C. T. & Short, C., 1879. *A Latin Dictionary.* Oxford: Clarendon Press.

Lietaer, B., 2002. *The Future Of Money.* London: Random House.

Löber, K. & Houben, A., 2018. Central bank digital currencies. *Bank for International Settlements.*

Lucas, R. E. J., 1977. Understanding Business Cycles. In: K. Brunner & A. H. Meltzer, eds. *Stabilization of the Domestic International Economy.* Amsterdam: North-Holland, pp. 7-29.

Lucas, R. E. J., 1980. Equilibrium in a Pure Currency Economy. *Economic Inquiry,* 18(2), pp. 203-220.

Lucas, R. E. J., 1982. Interest Rates and Currency Prices in a Two-Country World. *Journal of Monetary Economics,* 10(3), pp. 335-359.

Lucas, R. E. J., 1984. Money in a Theory of Finance. *Carnegie-Rochester Conference Series on Public Policy,* Volume 21, pp. 9-46.

Lucas, R. E. J., 1987a. *Models of Business Cycles.* New York: Basil Blackwell.

Lucas, R. E. J., 1987b. Money and Interest in a Cash-in-Advance Economy. *Econometrica,* 55(3), pp. 491-513.

Lucas, R. E. J. & Stokey, N. L., 1987. Money and Interest in a Cash-in-Advance Economy. *Econometrica,* pp. 491-513.

Macleod, H. D., 1866. *The Theory and Practice of Banking.* 2nd ed. London: Longmans, Green, Reader and Dyer.

Macleod, H. D., 1889. *The Theory of Credit.* London: Longmans, Green and Co..

Macleod, H. D., 1906. *The Theory and Practice of Banking.* London: Longmans, Green and Co..

Mandelbrot, B. B., 1983. *The Fractal Geometry of Nature.* New York: Henry Holt and Company.

Mankiw, N. G., 2013. *Macroeconomics.* 8th ed. New York: Worth Publishers.

Mattesini, F., Monnet, C. & Wright, R., 2009. Banking: A Mechanism Design Approach. *FRB of Philadelphia Working Paper No. 09-26.*

Matthews, K. & Thompson, J., 2005. *The Economics of Banking.* Chichester: John Wiley & Sons Ltd.

McCandless, G. T. J. & Weber, W. E., 1995. Some Monetary Facts. *Federal Reserve Bank of Minneapolis Quarterly Review,* 19(3), pp. 2-11.

McLeay, M., Radia, A. & Thomas, R., 2014a. Money in the modern economy: An introduction. *Bank of England Quarterly Bulletin,* Issue Q1, pp. 4-13.

McLeay, M., Radia, A. & Thomas, R., 2014b. Money creation in the modern economy. *Bank of England Quarterly Bulletin,* Issue Q1, pp. 14-27.

Merton, R. C., 1977. An analytic derivation of the cost of deposit insurance and loan guarantees: an application of modern option pricing theory. *Journal of Banking and Finance,* pp. 3-11.

Merton, R. C., 1993. Operation and Regulation in Financial Intermediation: A Functional Perspective. In: P. Englund, ed. *Operation and Regulation of Financial Markets.* Stockholm: Ekonomiska rådet, pp. 17-67.

Merton, R. C., 1995. A Functional Perspective of Financial Intermediation. *Financial Management,* 24(2), pp. 23-41.

Merton, R. C. & Bodie, Z., 1995. A Conceptual Framework for Analyzing the Financial Environment. In: *The Global Financial System: A Functional Perspective.* Boston: Harvard Business School Press, pp. 3-31.

Merton, R. C. & Bodie, Z., 2005. Design of Financial Systems: Towards a Synthesis of Function and Structure. *Journal of Investment Management,* 3(1), pp. 1-23.

Merton, R. C. & Thakor, R. T., 2018. Customers and Investors: A Framework for Understanding the Evolution of Financial Institutions. *Journal of Financial Intermediation.*

Mian, A. & Sufi, A., 2015. *House of Debt.* Chicago: The University of Chicago Press.

Minsky, H. P., 2008 [1986]. *Stabilizing an Unstable Economy.* New York: McGrawHill.

Mishkin, F. S., 1995. Symposium on the Monetary Transmission Mechanism. *The Journal of Economic Perspectives,* 9(4), pp. 3-10.

Mishkin, F. S. & Eakins, S. G., 2006. *Financial Markets and Institutions.* 5th ed. Boston: Pearson Education International.

Moore, B. J., 1988. *Horizontalists and Verticalists: The Macroeconomics of Credit Money.* Cambridge: Cambridge University Press.

Mouatt, S., 2011. Can Sequential and Non-Dualistic Interpretation of Marx's Monetary Thought be Synthesized with the Post-Keynesian Endogenous Money Paradigm?. *Critique: Journal of Socialist Theory,* 39(2).

Nason, J. M. & Cogley, T., 1994. Testing the implications of long-run neutrality for monetary business cycle models. *Journal of Applied Econometrics,* Volume 9, pp. S37-S40.

Pennacchi, G., 2012. Narrow Banking. *Annual Review of Financial Economics,* 4(0), pp. 1-36.

Phelan, G., 2016. Financial Intermediation, Leverage, and Macroeconomic Instability. *American Economic Journal: Macroeconomics,* 8(4), pp. 199-226.

Positive Money, 2014. *About Us.* [Online] Available at: http://www.positivemoney.org/about/

Revill, J., 2018. *Swiss voters reject campaign to radically alter banking system.* [Online] Available at: https://www.reuters.com/article/us-swiss-vote-sovereign/swiss-voters-reject-campaign-to-radically-alter-banking-system-idUSKBN1J60C0 [Accessed 10 June 2018].

Rey, H., 2018. Dilemma not Trilemma: The Global Financial Cycle and Monetary Policy Independence. *NBER Working Paper No. 21162.*

Rochon, L.-P., 1999a. *Credit, Money and Production: An Alternative Post-Keynesian Approach.* Cheltenham: Edward Elgar.

Rochon, L.-P., 1999b. The Creation and Circulation of Endogenous Money: A Circuit Dynamique Approach. *Journal of Economic Issues,* 33(1), pp. 1-21.

Rothschild, M. & Stiglitz, J., 1976. Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics,* 90(4), pp. 629-649.

Salyer, K. D., 1991. The Timing of Markets and Monetary Transfers in Cash-in-Advance Economies. *Economic Inquiry,* Volume XXIX, pp. 762-773.

Samuelson, P. A., 1958. An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money. *The Journal of Political Economy,* 66(6), pp. 467-482.

Santomero, A. M., 1984. Modeling the Banking Firm: A Survey. *Journal of Money, Credit and Banking,* Volume 6, pp. 576-602.

SARB, 2017. *Quartlery Bulletin (March),* Pretoria: South African Reserve Bank.

Scheffel, E. M., 2010. Business cycles, velocity and asset prices in a Wicksellian banking time economy. *PhD Thesis, Cardiff University.*

Schmandt-Besserat, D., 1980. The Envelopes That Bear the First Writing. *Technology and Culture,* 21(3), pp. 357-385.

Schmandt-Besserat, D., 1991. Two Precursors of Writing: Plain and Complex Tokens. In: W. M. Senner, ed. *The Origins of Writing.* s.l.:Univeristy of Nebraska Press, pp. 27-42.

Scholtens, B. & van Wensveen, D., 2000. A critique on the theory of financial intermediation. *Journal of Banking & Finance,* Volume 24, pp. 1243-1251.

Scholtens, B. & van Wensveen, D., 2003. The Theory of Financial Intermediation: An Essay On What It Does (Not) Explain. *Chapters in SUERF Studies,* Issue 1, pp. 7-53.

Schorfheide, F., 2000. Loss Function-Based Evaluation of DSGE Models. *Journal of Applied Econometrics,* Volume 15, pp. 645-670.

Schumpeter, J., 1939. *Business Cycles.* New York: McGraw-Hill Book Company, Inc..

Schumpeter, J. A., 1956. Money and the Social Product. *International Economics Papers,* Volume 6, pp. 148-211.

Schumpeter, J. A., 1983 [1934]. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle.* London: Transaction Publishers.

Schumpeter, J. A., 2006 [1954]. *History of economic analysis.* Abingdon: Taylor & Francis.

Semenova, A., 2011. Would You Barter with God? Why Holy Debts and Not Profane Markets Created Money. *The American Journal of Economics and Sociology,* 70(2), pp. 376-400.

Sheard, P., 2013. *Repeat After Me: Banks Cannot and Do Not "Lend Out" Reserves,* New York: Standard and Poor's Ratings Direct.

Shubik, M., 2016. Three Lectures on the Theory of Money and Financial Institutions: Lecture 1: A nontechnical overview. *Cowles Foundation Discussion Paper No. 2036.*

Sims, C. A., 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review,* 36(5), pp. 975-1000.

Sissoko, C., 2007. An Idealized View of Financial Intermediation. *Economics: The Open-Access, Open-Assessment E-Journal,* 1(5), p. 2007.

Sissoko, C., 2015a. The Bank-Centered View of the Money Market, Part I: Why Banks Are Different. *USC Gould School of Law Center for Law and Social Science Research Papers Series No. CLASS15-4.*

Sissoko, C., 2015b. The Bank-Centered View of the Money Market, Part II: Re-evaluating the Real Bills Approach to Macro-prudential Regulation. *University of Southern California Law School Legal Studies Working Paper Series No.158.*

Sissoko, C., 2016. Money is Debt. *Working paper,* May.

Solow, R. M., 1957. Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics,* 39(3), pp. 312-320.

Sowell, T., 2009. *The Housing Boom and Bust.* 1st ed. New York: Basic Books.

Spence, M., 1973. Job Market Signaling. *The Quarterly Journal of Economics,* 87(3), pp. 355-374.

Stein, J. C., 2013. Overheating in credit markets – origins, measurement, and policy responses. *BIS central bankers' speeches.*

Stiles, J., 2017. *Banks create money from nothing. And it gets worse.* [Online] Available at: https://thenewdaily.com.au/money/finance-news/2017/01/03/how-banks-create-money/ [Accessed August 2017].

Stock, J. H. & Watson, M. W., 1999. Business Cycle Fluctuations in US Macroeconomic Time Series. In: J. B. Taylor & M. Woodford, eds. *Handbook of Macroeconomics.* Amsterdam: North Holland, pp. 3-64.

Stockman, A. C., 1981. Anticipated inflation and the capital stock in a cash in-advance economy. *Journal of Monetary Economics,* 8(3), pp. 387-393.

Svensson, L. E. O., 1985. Money and Asset Prices in a Cash-in-Advance Economy. *Journal of Political Economy,* 93(51), pp. 919-944.

Taleb, N. N., 2012. *Antifragile: Things That Gain from Disorder.* New York: Random House Publishing Group.

Taylor, J. B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy,* Volume 39, pp. 195-214.

Taylor, J. B., 1996. How Should Monetary Policy Respond to Shocks While Maintaining Long-Run Price Stability? - Conceptual Issues. In: *Achieving Price Stability.* s.l.:Federal Reserve Bank of Kansas City, p. 181–195.

Thakor, A. V., 2014. Bank Capital and Financial Stability: An Economic Trade-Off or a Faustian Bargain?. *Annual Review of Financial Economics,* Volume 6, pp. 185-223.

The Libra Association, 2019. *Libra White Paper.* [Online]
Available at: https://libra.org/en-US/white-paper/ [Accessed 2019].

The New Economic Rights Alliance v Absa Bank Limited and others (2012).

Thornton, H., 1965 [1802]. *An Enquiry into the Nature and Effects of the Paper Credit of Great Britain.* New York: Augustus M. Kelley.

Tobin, J., 1956. The Interest-Elasticity of Transactions Demand For Cash. *The Review of Economics and Statistics,* 38(3), pp. 241-247.

Tobin, J., 1963. Commercial Banks as Creators of Money. *Cowles Foundation Discussion Paper No. 159.*

Tooke, T., 1844. *An Inquiry into the Currency Principle the Connection of the Currency with Prices, and the Expediency of a Separation of Issue from Banking,* London: Longman, Brown, Green, and Longmans.

Torchia, N. J., 1999. *Creatio Ex Nihilo and the Theology of St. Augustine: The Anti-Manichaean Polemic and Beyond.* New York, NY: Peter Lang Publishing Inc.

Townsend, R. & Wallace, N., 1987. Circulating Private Debt: An Example with a Coordination Problem. In: E. C. Prescott & N. Wallace, eds. *Contractual Arrangements for Intertemporal Trade.* Minneapolis: University of Minnesota Press, p. 176.

Trautwein, H.-M., 1996. Money, Equilibrium, and the Business Cycle: Hayek's Wicksellian Dichotomy. *History of Political Economy,* 28(1), pp. 27-55.

Turner, A., 2013. *Credit, Money and Leverage: What Wicksell, Hayek and Fisher know and modern macroeconomics forgot.* Stockholm: s.n.

Turner, A., 2015. *Between Debt and the Devil: Money, Credit, and Fixing Global Finance.* New Jersey: Princeton University Press.

von Mises, L., 1953 [1912]. *The Theory of Money and Credit.* New Haven: Yale University Press.

von Reden, S., 2007. *Money in Ptolemaic Egypt: From the Macedonian Conquest to the End of the Third Century BC.* Cambridge: Cambridge University Press.

Wallace, N., 1990. A Suggestion for Oversimplifying the Theory of Money. *Federal Reserve Bank of Minneapolis Quarterly Review,* 14(1), pp. 19-26.

Wallace, N., 2011. The Mechanism-Design Approach to Monetary Theory. In: B. M. Friedman & M. Woodford, eds. *Handbook of Monetary Economics.* Amsterdam: North-Holland, pp. 3-24.

Walsh, C. E., 2010. *Monetary Theory and Policy.* 3rd ed. Cambridge, Massachusetts: The MIT Press.

Wen, Y., 2005. Where's the Beef? The Trivial Dynamics of Real Business Cycle Models. *The Federal Reserve Bank of St. Louis Working Paper 2005-039A.*

Werner, R. A., 2014a. Can banks individually create money out of nothing? — The theories and the empirical evidence. *International Review of Financial Analysis,* Volume 36, pp. 1-19.

Werner, R. A., 2014b. How do banks create money, and why can other firms not do the same? An explanation for the coexistence of lending and deposit-taking. *International Review of Financial Analysis,* Volume 36, p. 71–77.

White, E. N., 1986. Before the Glass-Steagall Act: An Analysis of the Investment Banking Activities.of National Banks. *Explorations in Economic History,* Volume 23, pp. 33-55.

Wicksell, K., 1907. The Influence of the Rate of Interest on Prices. *The Economic Journal,* 17(66), pp. 213-220.

Wicksell, K., 1936 [1898]. *Interest and Prices: A study of the causes regulating the value of money.* London: Macmillan.

Wicksell, K., 1978 [1906]. *Lectures on Political Economy.* Fairfield, New Jersey: Augustus M. Kelley.

Williamson, S. & Wright, R., 2010. New Monetarist Economics: Methods. *Federal Reserve Bank of Minneapolis Research Department Staff Report 442,* April.

Williamson, S. & Wright, R., 2011. New Monetarist Economics: Models. In: B. M. Friedman & M. Woodford, eds. *Handbook of Monetary Economics.* Amsterdam: North-Holland, pp. 25-96.

Williams, Z., 2015. *You can print money, so long as it's not for the people.* [Online] Available at: http://www.theguardian.com/commentisfree/2015/oct/04/printing-money-jeremy-corbyn-quantitative-easing-peoples-qe

Wolf, M., 2014. *Strip private banks of their power to create money.* [Online] Available at: http://www.ft.com/intl/cms/s/0/7f000b18-ca44-11e3-bb92-00144feabdc0.html#axzz35pr1o1ps

Woodford, M., 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, New Jersey: Princeton University Press.

Wray, L. R., 2012. *Modern Money Theory: A Primer on Macroeconomics for Sovereign Monetary Systems.* Basingstoke, Hampshire: Palgrave Macmillan.

Wray, L. R., 2014. From the state theory of money to modern money theory: An alternative to economic orthodoxy. *Levy Economics Institute Working Paper No. 792.*

Zarlenga, S., 2002. *The Lost Science of Money.* Valatie: American Monetary Institute Charitable Trust.

Zelizer, V. A., 1989. The Social Meaning of Money: "Special Monies". *American Journal of Sociology,* September, 95(2), pp. 342-377.