Essays on Monetary Economics

Cristina Manea

TESI DOCTORAL UPF / Year 2020

THESIS SUPERVISOR
Jordi Galí and Alberto Martin
Department Departament d'Economia i Empresa

To my parents

Acknowledgements

I would like first to thank my parents for their constant support, as well as to my PhD advisors Jordi Galí and Alberto Martin for their exquisite academic guidance and advice while writing this thesis. A special thank you as well to Alain Durré who was a good friend, a mentor and *de facto* a third PhD advisor.

Another few (very generous) people have helped me pursue my studies and research in monetary economics over the years. They are, in chronological order, Véronique Joly (CMCO Strasbourg), Jalal El Ouardighi (Université de Strasbourg), Pierre Dehez (Université Catholique de Louvain), Olivier Pierrard (Université Catholique de Louvain, Banque Centrale de Luxembourg), Raf Wouters (National Bank of Belgium), Manfred Kremer (European Central Bank), Kirstin Hubrich (Federal Reserve Board), Boris Hofmann and Ryan Banerjee (both, Bank for International Settlements). Their support was crucial, and I am very grateful.

On the scientific side, my thesis has greatly benefited from the previous work of Professors Jordi Galí and Nobuhiro Kiyotaki in the fields of monetary economics and macrofinance. Working as a research assistant for three years for Professor Galí on the second edition of his textbook and on numerous academic papers allowed me to get a deep insight in his work and research methodoloy. Similarly, the intense well-structured PhD course by Professor Kiyotaki which I attended at the beginning of my doctoral program in Florence was a very rich hands-on learning experience. The academic inspiration has been further supplemented by a number of descriptive policy writings by Claudio Borio and Piti Disyatat (Bank for International Settlements) on the role of central banks nowadays and on how private banks create inside-money. I am filled with gratitude as well to many well-established researchers in monetary economics/macroeconomics whom I met at different stages of my doctoral program and who took the time to discuss research ideas with me. I hereby can name just a few, but they were many more: Pierpaolo Benigno, Florin Bilbie, Frederic Boissay, Fiorella de Fiore, Luca Fornaro, Simon Gilchrist, Peter Karadi, Narayana Kocherlakota, Giovanni Lombardo, Cyrill Monnet, Rigas Oikonomou, Louis Phaneuf, Huw Pill, Stephanie Schmitt-Grohe, Michael Woodford and Egon Zakrajsek.

I would like to extend my thanks more generally to the European Central Bank for hosting me for two years before my doctoral studies, first as intern in the research department, and then as a research analyst in Monetary Policy Strategy division. All three chapters of my PhD thesis were inspired by what I learnt during that period. A special thank you as well to the Bank for International Settlements (in particular to the Committee on Payments Systems, and the Monetary and Economic Department) for their well-written documentation on a wide range of topics related to central banking, and also for selecting me into their PhD Fellowship program last year. I also have the honor to thank for our recent discussion regarding the challenges faced by Eastern European countries in terms of monetary policy strategy to Mugur Isarescu, the Governor of the National Bank of Romania (my native country). My gratitude goes as well to the Federal Reserve Banks of Kansas City, of New York, of Richmond and of San Francisco, as well as to the Bank of Canada, the European Central Bank, Deutsche Bundesbank, Denmark National Bank (and more particularly, to Federico Ravenna), Banque Centrale de Luxembourg for inviting me for interviews in the past months. My papers have gained a lot from those interactions.

Furthermore, I would like to show appreciation to all my professors from my undergraduate and graduate studies in economics, as well as to participants to my seminars. Feedback on my papers from Andrea Caggese was particularly useful, as well as from Edouard Schaal and Davide Debortoli. The prompt advice I received from Raluca Vernic and Alexandru Bobe on a number of mathematical issues over the years was also invaluable. The preparation of my thesis benefitted also from constant support and advice on a wide range of administrative issues from Marta Araque and Laura Agustí Roig, and I would like to thank them with this occasion. I am grateful as well for the meetings and discussions I had with my collegues in the PhD program doing research in Monetary Economics (Christopher Evans, Christian Höynck, Mario Giarda, Angelo Gutierez, Derrick Kanngiesser), and for the good advice, encouragement and friendship of Stefanie Huber (former UPF PhD student of Professor Jordi Galí, now Assistant Professor at University of Amsterdam) and Flora Budianto (former PhD student Freie Universität Berlin, now economist at the Bank for International Settlements).

Last, but not least, I would like to thank as well to my family and close friends (Ioana Arsenie, Laura Margineanu, Lynn Pauly, Adela Panainte, Adriana Popescu– also my biology teacher from highschool) for their visits and unconditional support, and to Professor Costin Oancea for help with the Spanish translation of the abstract of my thesis.

Abstract

In my PhD thesis, I extend the basic New Keynesian (NK) model (Galí (2015), Woodford (2003)) on three distinct dimensions. (i) In the first chapter, I introduce endogenous money creation by private banks ("inside-money"). (ii) In the second chapter, I allow for a share of firms which face financial constraints, and I study how firm heterogeneity in terms of credit access affects monetary policy. (iii) In the third chapter, I analyze how the fiscal limit and the zero lower bound on the policy rate jointly constrain the optimal monetary-fiscal policy response to business cycle fluctuations. These extensions provide relevant insights for the ongoing review of monetary-policy frameworks.

Resum

Amplío el modelo básico basado en el Nuevo keynesianismo (Galí (2015), Woodford (2003)) en tres vertientes. (i) En el primer capítulo, introduzco la creación endógena de dinero por bancos privados. (ii) En el segundo capítulo, permito que una parte de las empresa pueda afrontar limitaciones financieras, y estudio cómo una heterogeneidad corporativa en relación al acceso crediticio afecta a la política monetaria. (iii) En el tercer capítulo, analizo cómo el límite fiscal y el nivel mínimo cero en la tasa de política monetaria, conjuntamente restringen la respuesta óptima de políticas monetaria y fiscal a las fluctuaciones cíclicas.

Contents

Introduction

Central banks around the world are now reviewing their monetary policy frameworks. Their current (flexible-)inflation targeting policies were designed based on the conclusions of the basic New Keynesian (NK) model (Galí (2015), Woodford (2003)). Those policy recommendations were derived abstracting away from features such as financial frictions, agent heterogeneity, the zero lower bound (ZLB) on the policy rate or fiscal limits. Advances in the literature however, as well as practice, have gradually hinted to the relevance of such features, and hence to a potentially necessary review of strategies used by central banks.

In my PhD thesis, I aim to contribute to the ongoing review of monetary policy frameworks on three different dimensions. In the first chapter, I introduce into the basic NK model endogenous money creation by the private banking sector (like deposits), or "inside money". I do so for two main reasons. First, I want to address concerns expressed by policymakers regarding the lack of explicit account of the monetary role of banks within this framework. I find that the new "inside money" model has the same equilibrium representation as the textbook "money-less" one, and hence transmission and optimal design of monetary policy in the two models are identical. Second, I want to create a standard benchmark with inside money issued by banks, which I plan to compare in the future to versions with alternative (global) forms of private money issued by non-banks such as the "Libra coin" recently announced by Facebook.

In the second chapter, I allow for a share of firms in the basic NK model to be financially-constrained. I find that the interactions between financially-constrained and unconstrained firms play a key role in the transmission of monetary policy. These interactions represent a new transmission channel which I call "the spillover channel". Because of this channel, the conclusions regarding the effect of financial frictions on the transmission and design of monetary policy in the heterogenous firm case may be very different than the weighted average of the predictions in the two polar cases (i.e. with only constrained, or only unconstrained firms). Another key insight from my analysis is that the ZLB affects differently constrained and unconstrained firms. Particularly, in economies where constrained firms normally respond more strongly to monetary policy than the unconstrained (which is likely the case of most advances economies nowadays), the ZLB particularly hurts constrained ones, while it benefits the unconstrained (conditional on a standard Taylor rule). For central banks, this implies that in the current low interest rate environment where the ZLB is occasionally binding, their policy is likely to have significant redistribution effects over the business cycle between the two types of firms.

Finally, in the third chapter, I explore, jointly with Alain Durré, how the proximity of the economy to its fiscal limit constrains the optimal monetary-fiscal policy response to business cycle fluctuations at the ZLB. We assume away both outright default on public debt and monetary financing. Our main result is that, as the economy approaches its fiscal limit, the reduction in fiscal space limits the future boom that the policymaker can promise at the ZLB, and hence, dynamics under optimal policy become less inflationary. The analysis is relevant in the current context of aging populations which simultaneously push many advanced economies (i) to their fiscal limit, and (ii) induce a decline in their long-run real interest rates (with the latter increasing the probability of hitting the ZLB). For central banks, this means that, if their economies keep converging to fiscal limit, they will gradually lose their ability to engineer concerted forward guidance policies with the fiscal authority so as to boost the economy at the ZLB.

Chapter 1

INSIDE-MONEY IN THE NEW KEYNESIAN MODEL

1.1 Abstract

The textbook New Keynesian framework has become a common tool for monetary policy analysis in central banks. Policymakers are nonetheless often concerned that this framework abstracts away from endogenous money creation, and lacks realism. To address this concern, I introduce endogenous money creation by the private banking sector (like deposits), or "inside money", into the textbook framework. I find that the new "inside money" model has the same equilibrium representation as the textbook "money-less" one, and hence, transmission and optimal design of monetary policy in the two models are identical[1].

Keywords: New Keynesian model, inside money, cashless, inside-liquidity banking theory

JEL Class.: E2 – E3 – E4

1.2 Introduction

A number of economists have expressed concerns over the lack of explicit account of banks' monetary role in New Keynesian (NK) models widely used for monetary policy analysis. These concerns have been expressed on two dimensions. On a first dimension, they regard the monetary fundamentals of cashless versions of these models, because they do not explicitly model the role of bank deposits ("inside-money") in transactions. For instance, Goodfriend and McCal-

[1] I thank my PhD advisor Jordi Gali and Piti Disyatat for comments.

lum (2007) consider the NK framework in Bernanke, Gertler and Gilchrist (1999) as "fundamentally non-monetary" because "it does not recognize the existence of a demand for money that serves to facilitate transactions". Similarly, Borio and Disyatat (2011) and Borio (2014) fear that model economies in such frameworks may fundamentally represent real (barter) ones, and that we may need better analytical representations of actual monetary economies.

On a second dimension, the concerns over the lack of explicit reference to the role of banks in money creation are related to the way banks are modeled in these setups (e.g. Ryan-Collins et al. (2011), Borio and Disyatat (2011), Borio (2014), Jakab and Kumhof (2015), Turner (2016)). Specifically, they are assumed to exclusively channel (real) resources from "savers" to "debtors"[2]. This hypothesis is however *a priori* at odds with the functioning of banking systems in practice where banks give loans by issuing deposits ("inside-money")[3].

I address these two related concerns in this paper. Regarding the first one, I find that the "cashless" basic NK setup (e.g. Woodford (2003), Galí (2015)) is isomorphic to monetary versions where bank deposits ("inside-money") issued within a perfectly competitive banking sector are used in transactions. Hence, this basic model is not inconsistent with the equilibrium dynamics of a monetary economy with inside money. Furthermore, to address the second dimension, I show that accounting explicitly for banks' monetary role in the canonical model of Bernanke, Gertler and Gilchrist (1999) is irrelevant for its equilibrium dynamics, namely that a version with banks giving loans by issuing deposits (in line with practice) is isomorphic to their original specification. These results do not imply however that banks' monetary role is generally irrelevant for monetary policy analyses within the NK paradigm. To make this point, I end with a number of research questions for which accounting for this role is essential.

To give a hint on the first result, note that according to the "cashless" definition in Woodford (1998), cashless models are meant to describe "pure-credit economies" with central bank liabilities playing the role of unit of account. Particularly, they are not meant to describe neither barter, nor monetary economies, but a third distinct category[4]. Specifically, according to Woodford (1998), the cashless setup describes a world where the execution of trade is decentralised and

[2]For instance, from households to entrepreneurs for investment purposes (Bernanke, Gertler and Gilchrist (1999)), or from patient households to impatient ones for consumption purposes (Curdia and Woodford (2016)).

[3]Detailed descriptions of the role of inside-money in modern payment systems are provided for example in publications by the Bank for International Settlements such as Committee on Payment and Settlement Systems (2003), Disyatat (2008) or Borio and Disyatat (2011).

[4]*Barter* is a simultaneous exchange of commodities, whether goods or labor services, with bargaining and without using money (The new Palgrave dictionary of economics (Vol. 1, pp. 384)) whereas in a *monetary economy* the medium of exchange is money (e.g. Collins Dictionary of Economics).

goods are diversified as in Lucas (1980), but where (nonbank) agents settle transactions on credit by issuing perfectly enforceable IOUs with all payments being carried out *via* book-keeping movements. Otherwise stated, the cashless economy is a hypothetical economy where there are no monetary frictions to justify the use in *transactions* of a distinct perfectly liquid asset such as "money" (Woodford (2003) p.31). Central bank liabilities ("outside-money") do however play a role as a "unit of account" in these setups. Specifically, the monetary authority sets the unit of account in terms of which prices (of both goods and financial assets) are quoted and controls the price level in the economy by setting the price of a one nominal unit of credit in the economy (by issuing a one period nominal bond).

Even though the mapping may not be explicit at first sight, the assumptions of these models are consistent with the role of central banks and the way transactions are settled nowadays in advanced economies. Specifically, central bank liabilities do play the role of unit of account[5]. Furthermore, even though (generally) nonfinancial agents cannot issue their own IOUs due to a lack of (multilateral) repayment commitment, bank deposits ("inside-money") play in practice the same role as the (underlying) IOUs issued in trade relations in the frictionless cashless version. Intuitively, in line with the actual functioning of banking systems, we may think of banks as being endowed with a repayment enforcement technology, and thus with the ability to exchange nonbank unenforceable IOUs ("bank loans") with their own enforceable IOUs ("bank deposits")[6].

The second result is more elaborate, but, as we will see, inherently linked to the first one. Hereafter, the chapter is organised as follows: Section 2 highlights some of the related theoretical literature; Section 3 describes the analytical setups and derives the two equivalence results; Section 4 discusses a number of research topics whose study within this framework requires an explicit account of banks' monetary role, whereas Section 5 concludes.

1.3 Contributions to the literature

The present analysis contributes to three main strands of literature. First, it lays out an extension of the NK paradigm. Particularly, it introduces for the first time in this setup bank deposits ("inside-money") created within a banking sector

[5]It is true that in practice central bank liabilities are also used as (outside-)money in transactions. It has been however shown that incorporating outside-money within the New-Keynesian setup has little quantitative significance (Woodford (2003), Ireland (2004), Woodford (2008)). The current analysis takes these results as given and focuses exclusively on the critique outlined in the introduction which concerns the lack of explicit reference to inside-money in such frameworks.

[6]For a description of the functioning of the banking sector in practice, see for instance Disyatat (2008), McLay et al. (2014 a-b), Werner (2014 a-b), Deutsche Bundesbank (2017).

modeled in line with the inside-liquidity theory proposed by Kiyotaki and Moore (2002) and it establishes an isomorphy to its cashless version. Most extensions of the NK model with "money" used in transactions identify it with non-interest bearing central bank liabilities (i.e. banknotes/coins)[7]. In few other extensions bank deposits are used in transactions (e.g. Stracca (2007), Goodfriend and McCallum (2008)), but banks' behaviour is not modelled in line with the inside-liquidity theory and these models don't feature the same neutrality result (in the absence of credit frictions and banks' operational costs)[8].

Second, the paper contributes to the cash-in-advance literature. Particularly, it shows that well-known monetary policy transmission channels such as the inflation/interest rate tax or the cost-channel which emerge when agents need to hold "cash" (i.e. liquid assets) in advance to pay for goods as in Lucas and Stokey (1987), and, respectively, when firms need to finance working capital before receiving proceeds on the sale of output as in Christiano and Eichenbaum (1992), vanish when (interest bearing) inside-money is used instead of (non-interest bearing) outside-money. Moreover, in the absence of nominal rigidities, the Friedman rule (zero nominal interest rates) no longer characterizes optimal monetary policy.

Third, the paper discusses the implementation of monetary policy via the interbank market in the context of the inside-liquidity banking theory developed by Kiyotaki and Moore (2002). This banking theory was developed within a "real" heterogenous agent macroeconomic framework without any reference to monetary policy.

1.4 The model

The analytical setup is based on the limit case of the monetary economy in Woodford (1998) where agents *do not extend any trade credit* among themselves due to a lack of trust. The setup is adjusted to be interpreted as a standard NK setup (as described for instance in Galí (2015) or Woodford (2003)) enriched with a liquidity(cash)-in-advance constraint. Nominal rigidities are abstracted away without any loss of generality.

To motivate the need for any monetary arrangement, as in Woodford (1998) which follows Lucas (1980, 1981), production and goods' exchanges are carried out in a "decentralised" fashion. Namely, firms are spatially scattered with workers selling labor to a particular firm (producing a particular variety) and consumers

[7] See for instance the derivation of the cashless limit in Woodford (1998).

[8] Disyatat (2011) models banks' behaviour in line with this theory (even though it does not explicitly refer to it), but the analytical setup is a partial equilibrium model used to study the implications of modelling banks' behaviour in this way (as opposed to the standard approach) for the bank lending channel of monetary policy transmission.

obliged to go to the location of each firm to buy the differentiated array of goods. At the beginning of each period, after shocks realise, markets open and equilibrium prices and quantities are determined. Goods' differentiation is a necessary condition for intra-period trade and the use of "money". If all goods were identical, each household would just consume the goods received as counterpart of its labor effort ("wage") and hence no exchange would take place ("autarchy"). However, since each worker wants to consume not only the good produced by the firm employing him, but instead a diversified basket of goods, trade emerges.

Barter would be one possible option as each worker may exchange the type of good produced at his firm with the goods received by other workers. As in Woodford (1998), I implicitly assume however that this option would entail large (unmodelled) costs making it unappealing in equilibrium[9]. Alternatively, I can assume that all exchanges must be made by the means of a perfectly multilaterally enforceable (liquid) asset. This role is played by central bank liabilities (banknotes, coins), namely "outside money" in standard NK models with monetary frictions. In the current analysis however, I consider the case of "inside-money". Specifically, I assume that even though households and firms cannot directly issue perfectly multilaterally enforceable IOUs, they can exchange them with the multilateral enforceable IOUs of private (trustworthy) third-party agents. The real counterpart of these third-party agents are "banks" and their functioning is modeled in line with inside-liquidity banking theory in Kiyotaki and Moore (2002).

The model economy is thus populated by a continuum of identical households, a continuum of (diversified) monopolistic firms, a perfectly competitive banking sector and a central bank. Banks and firms are owned by households.

1.4.1 Banks as inside-money suppliers

The behaviour of banks is modeled as in Kiyotaki and Moore (2002). Specifically, as already mentioned, the existence of banks is motivated by a lack of trust between private non-financial agents which prevents them from extending trade credit among themselves (i.e. issuing their own IOUs to purchase goods/services on credit). Banks are assumed to have a "multilateral commitment" technology and to exchange their own "multilateral enforceable" IOUs ("bank deposits") for the "unenforceable" IOUs of a nonfinancial agent ("bank loan") who needs to buy on credit (figure 1.1).

[9]Note that, if all agents perfectly trusted each other, another option would be for consumers to issue units of credit (IOUs) to purchase goods from each firm, and at the end of each period, these units of credit to be settled among firms and workers (given the return on labor to which they are entitled). In this case an additional perfectly liquid asset such as "money" would play no special role in the economy. This is the very meaning of the cashless limit. As agents do not trust each other however this option is not implementable anymore in the current context.

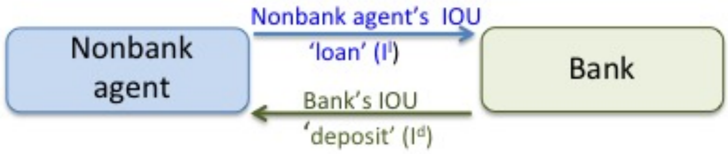


Figure 1.1: Banks' role in the model in line with inside-liquidity theory

Since banks are trust-worthy in the sense that their liabilities are perfectly enforceable, the seller of the "credit good" accepts them whereas it would not accept the IOUs of its direct (nonbank) trade partner. Operational costs of banks are normalised to zero. Furthermore, since banks are trusted by all agents in the economy, their liabilities are multilaterally enforceable (not only bilaterally), namely they are perfectly liquid, and thus they can circulate as "money". Without loss of generality, all IOUs issued in the economy are one-period IOUs.

The banking sector is composed by a large number of identical banks interacting on three different perfectly competitive markets: loan, deposit and interbank markets. Equilibrium in the loan market determines the one-period (loan) interest rate that non-financial agents need to pay to banks, whereas equilibrium in the deposit market determines the one period (deposit) interest rate that banks need to pay to (non-financial) bearers of their liabilities issued as a counterpart of loans[10]. Deposits issued by all banks are identical and thus they can be exchanged at par value. Whenever this happens the initial bank issuer of the IOU enters a credit relation with the new bank (figure 1.2). The interest rate on such "interbank credit" is decided on the interbank market. It is assumed without loss of generality that private credit can only be intermediated via the private banking sector (namely, households and firms cannot exchange their IOUs, or banks' IOUs directly with the ones of the central bank).

1.4.2 Monetary policy implementation with inside-money

In the cashless basic NK model (Galí (2015), Woodford (2003)), monetary policy controls nominal (and real) interest rates in the economy by setting the price of a one-period government bond in zero net supply. In the "inside-money" economy, the central bank controls the one-period nominal interest rate in the economy by controlling the interbank market rate. It does so by committing (i) to exchange its

[10]Note that the bearers of banks' IOUs may change within the period as deposits are used in transactions. For instance, when households take a bank loan to finance consumption, the initial bearer of the paper is a household, then the IOU is transferred to a firm and then it returns to a household once wages are paid at the end of the period.

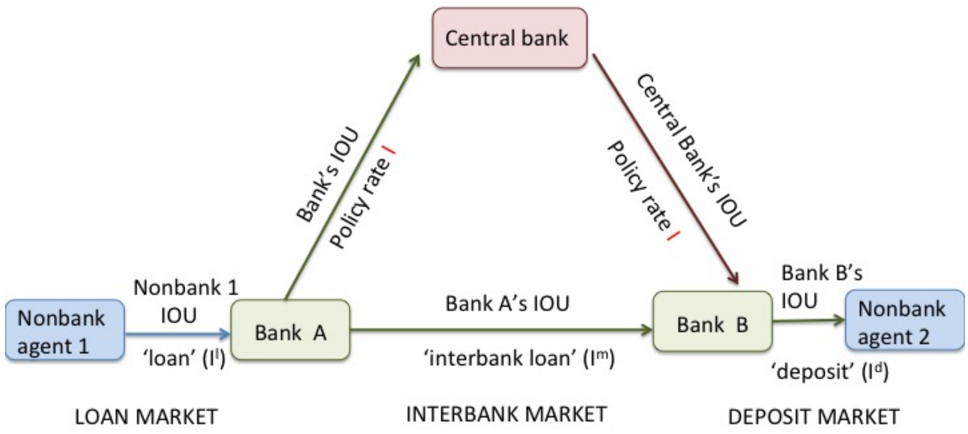


Figure 1.2: Monetary policy implementation *via* the interbank market

own IOUs with the IOUs issued by any bank in the economy, and to pay a chosen interest rate ("policy rate") on its IOUs ("deposit facility"), and (ii) to require the same interest rate from the original bank issuer of the IOUs ("borrowing facility"). At equilibrium, no bank has an incentive to require for an interbank market loan an interest rate higher than the policy rate. This can be proved by contradiction. Assume the creditor bank (Bank B in figure 1.2) asked for an interbank rate higher than the policy rate[11]. Then the debtor bank (Bank A in figure 1.2) could always exchange the IOU of Bank B for central bank's IOUs and pay instead the (lower) policy rate. In this case Bank B (the creditor bank) would necessarily receive the (lower) policy rate from the central bank[12]. Note that banks in this setup are indifferent between extending credit directly among themselves at the policy rate or making use of the deposit and borrowing facilities of the central bank.

Furthermore, no bank has an incentive to pay an interest on deposits different than the policy rate. If a bank paid a lower deposit rate, then nonbank agents would exchange them for the IOUs (deposits) of another bank. In this case, the initial bank would have to (eventually) pay to the new bank the interbank market rate which is equal to the policy rate. Thus, at equilibrium, all banks necessarily

[11]Note that the interbank market rate cannot be lower than the policy rate. If this were true, the creditor bank would not lend funds directly to the debtor bank and would prefer to exchange instead the IOU of the debtor bank with the one of the central bank to gain a higher interest rate (the policy rate).

[12]Note how the implementation is in line with the actual functioning of the banking sector: with cash holdings set to zero, if certain banks choose to use the "deposit facility" of the central bank instead of lending their overnight deposits surplus to other banks on the interbank market, other banks in the system will necessarily have to refinance themselves at the central bank via the "borrowing facility".

set the deposit rate equal to the policy rate.

And finally, the loan interest rate equals also at equilibrium the policy rate. If a bank required a loan interest rate higher than the deposit rate (which equals the policy rate), then demand for its services would be zero because there would always exist other banks which can propose one period loans at the (strictly lower) policy rate. Thus, at equilibrium, loan, deposit and interbank markets all clear at a nominal interest rate i_t equal to the policy rate chosen by the central bank.

1.4.3 A cashless economy *versus* a monetary economy with inside-money

I now tackle the first concern over the lack of monetary fundamentals of cashless NK models by establishing an isomorphism with their monetary versions with liquidity-in-advance constraints and inside-money (bank deposits). In the version presented in this section, households receive their wage at the end of each period, but need to consume at the beginning of the period (households face "cash-in-advance" constraints). In the second version, included in the Appendix on page 22, firms need to finance the wage bill at the beginning of the period before receiving proceeds on sales ("working-capital-in-advance" constraint). Importantly, relatively to the cashless basic NK model, in the "inside-money" versions there is no government bond in zero-net-supply. There is no need for this asset since monetary policy is implemented via the interbank market. Private agents can save instead by investing in one-period bank deposits.

In the first version where the representative household receives the wage at the end of each period, its behaviour is described following the cash-in-advance literature by considering separately a liquidity-in-advance constraint (1.1), an equation showing its outstanding wealth at the end of the period (1.2), and a solvency condition (1.3). The liquidity-in-advance constraint (1.1) states that *at the beginning of the period* households can pay for consumption goods C_t at price P_t in two ways. One is by using maturing bank deposits $(1 + i_{t-1})D_{t-1}$. The other is by issuing debt to firms via the banking sector, namely by exchanging their own IOUs (which are not enforceable, and hence cannot be used to buy goods on credit) with the ones of banks (which are multilaterally enforceable), and by transferring the latter to firms (as in figure 1.1).

The second option results in households having a liability towards banks ("households taking a bank loan"), and, banks, in turn, having a liability towards firms ("firms receiving banks' deposits"). Following convention in the literature, the liabilities of banks towards non-bank agents ("bank deposits") are denoted by $D_t \geq 0$. Thus, $D_t < 0$ stands for the liabilities of a non-bank agent towards banks ("bank loan"). I distinguish between *new* liabilities of banks towards nonbank agents *created* at the beginning of the period which I denote by D_t^*, and *outstanding* liabilities of banks towards non-bank agents at the beginning of the period which I denote by D_t. For the household, $-D_t^*$ denotes IOUs issued by households to banks at the beginning of period t ("bank loans" contracted at the beginning of the period), whereas D_t denotes outstanding IOUs of banks towards households at the end of period t (outstanding "bank deposits" at the end of t).

The second constraint (1.2) states that the value of nominal wealth held as bank deposits at the end of the period equals the wage income $W_t L_t$ (transferred in the form of deposits at the end of period, with W_t the nominal wage rate and L_t the number of labor units), firms' dividends Div_t, the initial amount of deposits $\Big((1+i_{t-1})D_{t-1} + (-D_t^*)\Big)$, net of consumption expenditures $P_t C_t$ and (outstanding) bank credit contracted at the beginning of the period $(-D_t^*)$. Specifically, each household solves

$$\max_{C_t, L_t, D_t, D_t^*} E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \qquad \text{subject to:}$$

$$\mu_t : P_t C_t \leq (1 + i_{t-1})D_{t-1} + (-D_t^*) \tag{1.1}$$

$$\lambda_t : D_t = W_t L_t + Div_t + \Big((1 + i_{t-1})D_{t-1} + (-D_t^*)\Big) - P_t C_t - (-D_t^*) \tag{1.2}$$

$$\lim_{T\to\infty} E_t \Big\{ \beta^{T-t} \frac{U_{c,T}}{U_{c,t}} \frac{D_T}{P_T} \Big\} \geq 0 \tag{1.3}$$

where C_t is a standard Dixit-Stiglitz consumption index of a continuum of varieties i indexed on the unit interval $C_t \equiv \Big(\int_0^1 C_t(i)^{1-\frac{1}{\epsilon}} di \Big)^{\frac{\epsilon}{\epsilon-1}}$ with the associated price $P_t \equiv \Big(\int_0^1 P_t(i)^{1-\epsilon} di \Big)^{\frac{1}{1-\epsilon}}$. The first-order Kuhn-Tucker conditions for a maximum are:

$$\begin{aligned}
C_t : U_{c,t} + \mu_t P_t + \lambda_t P_t &= 0, \\
L_t : U_{l,t} - \lambda_t W_t &= 0, \\
D_t : -E_t\big\{\beta \mu_{t+1}\big\}(1 + i_t) + \lambda_t - E_t\big\{\beta \lambda_{t+1}\big\}(1 + i_t) &= 0, \\
D_t^* : \mu_t &= 0, \\
\mu_t : \Big((1 + i_{t-1})D_{t-1} - D_t^* - P_t C_t\Big)\mu_t &= 0, \\
\mu_t \leq 0, \ \ (1 + i_{t-1})D_{t-1} - D_t^* - P_t C_t &\geq 0, \\
\lim_{T\to\infty} E_t \Big\{ \beta^{T-t} \frac{U_{c,T}}{U_{c,t}} \frac{D_T}{P_T} \Big\} &= 0,
\end{aligned}$$

and the equation showing its outstanding wealth at the end of the period (1.2).

Note that in this setup the liquidity-in-advance constraint is always slack ($\mu_t = 0$) i.e. it does not constrain household's choice. The reason is that in the 'inside-money' economy households can costlessly spend in advance their period wage income by means of bank credit. Thus, after combining previous equations, we can describe the behaviour of the representative household using the same equations as in the cashless basic NK model (i.e. standard Euler, labor supply and household budget constraint equations):

$$\beta(1+i_t)E_t\Big\{\frac{U_{c,t+1}}{U_{c,t}}\frac{P_t}{P_{t+1}}\Big\} = 1,$$

$$-\frac{U_{l,t}}{U_{c,t}} = \frac{W_t}{P_t},$$

$$W_tL_t + Div_t + (1+i_{t-1})D_{t-1} = P_tC_t + D_t.$$

The supply-side of the economy is also identical to the one in the flexible price version of the basic cashless NK model (e.g. Galí (2015), Chapter 3). Specifically, it is composed by a continuum of diversified firms in monopolistic competition with a Cobb-Douglas production technology. The input market (here, labor market) is competitive and firms act as price-takers. Firms need to pay inputs (here, wages) by the means of a perfectly liquid financial asset. They cannot issue such an asset, but they can use banks IOUs (banks' deposits) for this purpose (either by exchanging its own IOUs with banks' IOUs, or by using outstanding holdings of such assets). Since in this version firms pay wages at the end of the period, and sell their products within the period, they face no binding liquidity- constraints. Specifically, the IOUs received from selling goods are used to pay workers and shareholders. Thus, the problem of firm i writes

$$\max_{P_t(i),Y_t(i),L_t(i)} P_t(i)Y_t(i) - W_tL_t(i) \qquad \text{subject to:}$$

$$Y_t(i) = A_tL_t(i)^{1-\alpha}$$

$$Y_t(i) = \Big(\frac{P_t(i)}{P_t}\Big)^{-\epsilon}Y_t$$

where $Y_t(i)$ stands for output of firm i and A_t is an exogenous productivity process. The behaviour of firm i is thus described by

$$P_t(i) = \mathcal{M}\frac{W_t}{(1-\alpha)A_tL_t^{-\alpha}(i)}$$

$$Y_t(i) = A_tL_t(i)^{1-\alpha}$$

$$Y_t(i) = \Big(\frac{P_t(i)}{P_t}\Big)^{-\epsilon}Y_t$$

Standard market clearing conditions apply on goods and labor markets. At equilibrium, from representative household's budget constraints (1.1) and (1.2), we get $D_t = (1 + i_{t-1})D_{t-1}$. For $D_{-1} = 0$, this implies $D_t = 0, \ \forall t$. Finally, nominal determinacy is achieved, for instance, by having the central bank follow a policy rule which ensures equilibrium uniqueness.

We can thus conclude that the equations describing the dynamics of (real or nominal) variables in the model are identical to the ones in the cashless basic NK model (here, its flexible price version) despite the liquidity-in-advance constraints of private non-financial agents motivated by the decentralization of trade, and the lack of trust between these agents (i.e. lack of reenforceability of trade credit among them). Otherwise stated, the monetary version with inside-money issued within a perfectly competitive banking sector is isomorphic to the cashless one. Furthermore, I show in the Appendix 1.8.1 on page 22 that if firms need to pay instead wages at the beginning of the period in advance of sales (i.e. they have a"working capital in advance constraint"), the same isomorphism between the version with inside-money and the cashless version emerges.

1.4.4 Inside-liquidity banking theory and NK models

In this section I tackle the second concern regarding the way banks are modeled when explicitly included in NK models, namely the lack of reference to their role in money creation. I take as a reference the model in Bernanke, Gertler and Gilchrist (1999) (hereafter, BGG (1999)), and I compare the equilibrium dynamics of the original model where financial intermediaries take deposits from households and lend them to entrepreneurs for investment purposes, with a version where they behave in line with the inside-liquidity banking theory. The choice of this model is without loss of generality. Even when financial intermediaries are explicitly identified with "banks" (e.g. Gertler and Kiyotaki (2011), Gertler and Karadi (2011)), their behaviour is described in the same manner as in BGG (1999). To focus strictly on the impact of the two different approaches to modeling banks' behaviour on equilibrium dynamics, I abstract from financial frictions and nominal price rigidities.

The setup is an extension of the one described in the previous section with production run by perfectly competitive risk-neutral entrepreneurs who use both labor and physical capital as inputs, and finance physical capital both with their own funds (retained earnings) and loans from a financial intermediary.

Non-financial agents Each period, entrepreneurs make two types of choices: a production decision given their outstanding physical capital, and a capital investment decision for production in the following period. As in BGG (1999), firms resell and rebuy each period their entire capital stock on the market. Namely, given capital level K_t chosen in the previous period, each entrepreneur i solves

$$\max_{Y_t(i),L_t(i)} P_t(i)Y_t(i) - W_tL_t(i) \qquad \text{subject to:}$$
$$Y_t(i) = A_tK_t(i)^\alpha L_t(i)^{1-\alpha}$$

The production technology is assumed to exhibit constant returns to scale. Production choices by each entrepreneur i given its capital stock are described by

$$P_t(i) = \frac{W_t}{(1-\alpha)\frac{Y_t(i)}{L_t(i)}},$$
$$Y_t(i) = A_tK_t(i)^\alpha L_t(i)^{1-\alpha}.$$

Importantly, note that in contrast to the model in the previous section, sales income now comes from *both* households (consumption) and firms (investment).

Following BGG (1999), capital investment decisions are taken at the end of each period (given expected return), and risk neutral entrepreneurs are willing to absorb the associated macroeconomic risk. Capital is homogenous, namely newly produced capital units within the period have the same value as older vintages and thus sell at the same price Q_t.

The whole stock of capital is financed each period by bank credit repaid with interest in the following period (short term debt) and entrepreneurial net worth (to be defined latter). Entrepreneurs invest in capital goods until the expected nominal return $E_t\{R^k_{t+1}\}$ equals the gross nominal loan interest rate. The latter equals the one-period gross nominal interest rate $1+i_t$ since aggregate risk is borne by firms and thus loans are risk-free for banks[13]:

$$E_t\{R^k_{t+1}\} \equiv E_t\left\{\frac{\frac{\alpha P_{t+1}(i)Y_{t+1}(i)}{\mathcal{M}K_{t+1}(i)} + Q_{t+1}(1-\delta)}{Q_t}\right\} = 1+i_t$$

where δ denotes the depreciation rate and $\frac{\alpha P_{t+1}(i)Y_{t+1}(i)}{\mathcal{M}K_{t+1}(i)}$ is the expected marginal (nominal) return of capital.

As in BGG (1999), entrepreneurs have finite lives and a constant survival probability γ to the next period. The birth rate of entrepreneurs is such that the fraction of agents who are entrepreneurs is constant. This assumption avoids the case where the entrepreneurial sector ultimately accumulates enough wealth to be fully self-financing. Entrepreneurs dying in period t are not allowed to invest in capital, but instead simply consume their retained earnings. Furthermore, total labor input

[13] I assume that the return to capital is sensitive only to aggregate risk since idiosyncratic risk does not play any particular role in the absence of financial intermediation frictions.

L_t is a composite index of household labor L_t^h, and "entrepreneurial labor", L_t^e, namely $L_t = (L_t^h)^{\Omega}(L_t^e)^{1-\Omega}$, with entrepreneurial labor supplied inelastically and total entrepreneurial labor normalized to unity. Entrepreneurial labor is used to make new firms start with some initial net worth. End-of-period net worth of all entrepreneurs surviving to the next period equals

$$N_{t+1} = \gamma\Big(Q_tK_t - \big(1+i_{t-1}\big)\big(Q_{t-1}K_t - N_t\big) - \delta K_tQ_t\Big) + W_t + \big(P_tY_t - W_tL_t\big) \quad (1.4)$$

where Q_tK_t is the market value of outstanding capital holdings, $(1+i_{t-1})(Q_{t-1}K_t - N_t)$ is the gross nominal interest rate paid on the loan taken in the previous period to finance capital, δK_tQ_t is the market value of depreciated capital goods, W_t is the entrepreneurial wage. Since firms are now owned by entrepreneurs, households receive no dividends (i.e. $Div_t = 0$). The aggregate physical capital dynamics are described by[14]:

$$K_{t+1} = (1-\delta)K_t + \mathcal{I}_t \quad (1.5)$$

where $\mathcal{I}_t$ is net investment in period t, and δ is the capital depreciation rate.

Households behave exactly as in the model used in the previous section.

"Banks" – mainstream versus inside-liquidity theory In mainstream macroeconomic models such as BGG (1999) savings of households are channelled through financial intermediaries to fund the acquisition of physical capital goods by firms[15]. Even when financial intermediaries are explicitly identified with banks, their behaviour is described in the same manner (e.g. Gertler and Kiyotaki (2011), Gertler and Karadi (2011))[16]. As shown next however, this is without loss of generality since when the (more realistic) inside-liquidity theory of banking is applied instead, despite its distinct narrative, equilibrium dynamics of resulting models are identical.

Specifically, consider the case of firms willing to purchase capital goods in the inside-liquidity banking theory case. The potential sellers of such goods are other firms in the economy, namely the producers of new capital goods and owners of older capital vintages willing to liquidate them. If firms trusted each other, they would extend trade credit directly among themselves i.e. they would issue

[14]BGG (1999) additionally included increasing marginal adjustment costs in the production of capital to allow a variable price of capital. To ease exposition, I abstract from such costs since they are not relevant for the current argument.

[15]BGG (1999) p.1349: 'The entrepreneur borrows from a financial intermediary that obtains its funds from households'.

[16]Sometimes banks' are modelled as facing themselves frictions in financing their loan portfolio and thus their own net worth may impact real economic dynamics (e.g. Gertler and Kiyotaki (2011), Gertler and Karadi (2011)).

one period IOUs to the seller of capital goods equal to the value of capital goods augmented interest (equal in equilibrium to the expected capital return)[17].

However, since this is not the case, in each such trade credit relation a bank acts as an intermediary by exchanging its own one period IOUs ('bank deposits') for the one period IOUs of the 'buyer firm' ('loan'), banks' IOUs being the ones given to the 'seller firm' in exchange for capital goods (figure 1.1). As a result entrepreneurs will issue (among themselves) an aggregate amount of bank deposits equal to the market value of the fraction of capital goods externally financed, namely $\Big(Q_t K_{t+1} - N_{t+1}\Big)$. In the aggregate, these deposits are received as (current) sale revenue and are used both to pay (current) wages and, as part of internal funds, to buy investment goods.

Combing the representative household budget constraint (1.2) with the goods market clearing condition $Y_t = C_t + C_t^e + \mathcal{I}_t$ and setting $Div_t = 0$, we obtain that the following relation is satisfied in equilibrium:

$$W_t L_t^h + (1 + i_{t-1}) D_{t-1} = P_t Y_t - P_t C_t^e - P_t \mathcal{I}_t + D_t. \tag{1.6}$$

By further replacing the expression of $(P_t Y_t - W_t L_t^h)$ from the equation above in the expression of aggregate net worth of entrepreneurs (1.4), and using $W_t L_t = W_t L_t^h + W_t$, it yields:

$$\begin{aligned}
N_{t+1} &= \gamma\Big(Q_t K_t - (1 + i_{t-1})\big(Q_{t-1} K_t - N_t\big) - \delta K_t Q_t\Big) + W_t + (1 + i_{t-1}) D_{t-1} \\
&\quad + P_t C_t^e + P_t \mathcal{I}_t - D_t - W_t \\
&= \gamma\Big(Q_t K_t - (1 + i_{t-1})\big(Q_{t-1} K_t - N_t\big) - \delta K_t Q_t\Big) + (1 + i_{t-1}) D_{t-1} + \\
&\quad + P_t C_t^e + P_t \mathcal{I}_t - D_t
\end{aligned}$$

After replacing the expression of aggregate entrepreneurial consumption $P_t C_t^e = (1-\gamma)\Big(Q_t K_t - (1+i_{t-1})\big(Q_{t-1} K_t - N_t\big) - \delta K_t Q_t\Big)$, the relation above becomes:

$$\begin{aligned}
N_{t+1} &= \Big(Q_t K_t - (1 + i_{t-1})\big(Q_{t-1} K_t - N_t\big) - \delta K_t Q_t\Big) + (1 + i_{t-1}) D_{t-1} + \\
&+ P_t \mathcal{I}_t - D_t
\end{aligned}$$

Further using the expression of net capital investment $\mathcal{I}_t$ from the law of motion of capital (1.5) and that $Q_t = P_t$ yields:

$$\begin{aligned}
(Q_t K_{t+1} - N_{t+1}) &= (1 + i_{t-1})\big(Q_{t-1} K_t - N_t\big) - (1 + i_{t-1}) D_{t-1} + D_t \\
(Q_t K_{t+1} - N_{t+1} - D_t) &= (1 + i_{t-1})\big(Q_{t-1} K_t - N_t - D_{t-1}\big)
\end{aligned}$$

[17] In BGG (1999), entrepreneurs are risk neutral, and hence willing to absorb the macroeconomic risk from capital investment.

For $Q_{t-1}K_t - N_t - D_{t-1} = 0$, this implies

$$Q_tK_{t+1} - N_{t+1} = D_t, \ \ \forall t$$

or, the value of capital financed by entrepreneurs via the banking sector, equals in equilibrium the value of one period deposits held by households.

Thus, when applying inside-liquidity banking theory, end-of-period savings of households D_t (in terms of bank deposits) also equate in equilibrium the market value of investment externally financed by entrepreneurs via the banking sector $Q_tK_{t+1} - N_{t+1}$. So, even though the two approaches to the role of banks in the economy have distinct narratives, these narratives imply isomorphic equilibrium dynamics[18].

We may thereby conclude that *under the assumptions of mainstream macroeconomic models*, where the liquid nature of banks' liabilities does not play any specific role, the explicit monetary role of banks is irrelevant for equilibrium dynamics. Thus, the abstraction made is without loss of generality. For completeness, the equations describing the equilibrium dynamics of the model specification with capital investment encompassing the narratives of both banking theories are summarized in the Appendix 1.8.2 on page 23.

1.5 On the relevance of banks' monetary role

These equivalence results do not imply however that the monetary role of banks should generally be thought as irrelevant for monetary policy within the NK paradigm. To make this point, I discuss in this section some cases where modeling it explicitly, as we do in our current analysis, is consequential.

To begin with, the inside-money version of the basic NK model presented in section 1.4.3 could serve as an analytical exposition for the cash abolishment proposal made by Rogoff (2017) to eliminate the zero lower bound constraint (and associated inefficiencies)[19]. Specifically, it allows one to get an intuition on how monetary policy could be implemented if the central bank stopped issuing banknotes and coins and why the zero lower bound would become irrelevant in such a world[20]. Furthermore, it helps highlighting how negative nominal interest rates are only a convention strictly related to the unit of account and there is no

[18]In particular, note that following the inside-liquidity narrative, it is the gross investment level (i.e. the value of capital in the economy) which determines the end of period level of outstanding savings, whereas it is the other way round in the alternative (conventional) case.

[19]I thank my PhD co-advisor Jordi Galí for bringing to my attention the existence of this book and of a potential link to the analysis in my paper.

[20]Note that in the current setup with inside-liquidity created within the banking sector monetary policy faces no constraints in setting the policy rate to negative values.

conceptual difference between positive and negative nominal interest rates. In the words of Rogoff (2017), negative rate policy would just be "central banking business as usual, namely cutting interest rates in negative territory would work the same way as interest rate cuts in positive territory".

Importantly, the inside-money versions described in this paper provide a first explicit modeling within the NK framework of how a central bank could implement its policy if it stopped supplying *non-interest liabilities*. Standard cashless setups feature a zero-lower bound (ZLB) constraint because, as long as policy rates are strictly positive, such liabilities ("cash") are supplied by the central bank even though they are not held in equilibrium (e.g. Woodford (1999), Woodford (2003), Chapter 1)[21]. Otherwise stated, the term "cashless" strictly refers to the equilibrium outcome when nominal interest rates are strictly positive. Once the policy rate reaches the ZLB however, the central bank loses in these standard setups its power to influence nominal (and real) interest rates in the economy i.e. the model economy enters a "liquidity trap".

Second, the explicit incorporation of "inside-money" in the NK framework makes clear that interpreting "money" in extensions with "money-in-the-utility" (MIU) differently than "monetary base (M0)", namely M1, M2 or M3 (e.g. M1 in Cooley and Hansen (1989, 1991), M2 in Ireland (2004)), is problematic. Note that in the basic NK model with inside-money (both its cash-in-advance and working capital-in-advance versions) agents use in transactions exclusively monetary aggregates whose real counterpart is (the privately issued fraction of) M1. Nevertheless, the setup is isomorphic to the "cashless" version of the model (which

[21](i) Woodford 1999 page 34: "note that the equations [of the cashless model] are not simply the cashless limit of the equilibrium conditions of a monetary economy; they are also the equilibrium conditions that must be satisfied by the real interest rate and real financial wealth in a completely non-monetary economy; thus they could easily be derived by abstractly entirely from the use of money in transactions. *The only reason that I have described the system consisting of these equations together with the policy rules as determining the price level in the cashless limit of a monetary economy- rather than simply saying that they describe price level determination in an economy where cash is not needed for transactions is that it is not clear that a central bank should have any way of implementing the policy rule when money is not used at all, even though it can implement such a rule in a monetary economy no matter how close to zero alfa may be.*" (ii) Woodford 1999 page 26: "the central bank controls the rate on the market for short term nominal debt by staying ready to exchange public debt for money in arbitrary quantities at the price that it has decided upon. It is not possible for the central bank to bring about an interest rate $R < 1$ since this would be inconsistent with equilibrium owing to the arbitrage opportunity that it would create". (iii) Woodford 2003 Chapter 1.2. page 75: "The function [of the interest rate target rule] is assumed to be nonnegative on the grounds that it is not possible for the central bank to drive nominal interest rates to negative levels. I assume that, as under typical current arrangements, the holders of central bank balances have the right to ask for currency in exchange for such balances at any time and that it is infeasible to pay negative interest on currency. Hence an attempt to pay negative interest on central bank balances would lead to zero demand for such balances and a market overnight interest rate of zero rather than a negative overnight interest rate".

is not the case for MIU specifications). Same logic would apply for M2 and M3. Thus, as pointed out by Woodford (2003) pp. 117, "money" should be strictly interpreted in these frameworks as central bank liabilities in positive net supply (i.e. the monetary base M0, "outside-money" as opposed to "inside-money").

Third, and most importantly, one could imagine NK setups populated by several types of financial intermediaries, where the high degree of multilateral enforceability of banks' liabilities is relevant for the transmission of monetary policy and for macroeconomic dynamics and allocation. Such analyses may uncover insights for instance regarding differences in monetary policy transmission and in the responses of macroeconomic variables to shocks at business cycle frequencies for economies with different financial structures such as the bank-based case in the Euro Area and the market-based one in the United States[22]. One promising starting point to build such an extension could be the model in Kiyotaki and Moore (2018) which explicitly takes into account that "fiat money" issued by the central bank is more liquid than equity. Specifically, one could replace "fiat outside money" with "bank inside-money", and add in the analysis frictions specific to financial intermediation via the banking sector.

1.6 Conclusions

I addressed in this paper two related concerns expressed by economists working at monetary policy institutions regarding the lack of explicit account of banks' monetary role in NK models, namely (i) the lack of reference to bank deposits in cashless versions of these models, and (ii) the lack of explicit account of banks' monetary role once included in this framework. I tackled the first concern by showing that the cashless specification is isomorphic to monetary versions with bank deposits ('inside-money') used in transactions, and the second one, by showing that banks' monetary role is irrelevant for equilibrium dynamics under the assumptions of standard models. In the end, with the help of some examples, I pointed out however that these results do not imply that banks' monetary role should generally be thought as irrelevant for monetary policy analyses within this paradigm. To my knowledge, there are no extensions of the NK framework where banks behaviour is modeled in line with inside-liquidity theory and banks' role in the supply of liquidity is relevant for the equilibrium allocation. Exploring this line of research may however uncover new interesting findings regarding the transmission and optimal design of monetary policy.

[22]This research idea was first brought to my attention long before writing this paper during the time I was working as a research analyst in the Monetary Policy Strategy division of the European Central Bank by Jens Eisenschmidt, Principle Economist in the division.

1.7 References

Barth III, Marvin J., and Valerie A. Ramey (2001): ”The cost channel of monetary transmission,” NBER macroeconomics annual, 16, 199-240.

Borio, Claudio (2014): ”The financial cycle and macroeconomics: What have we learnt?,” *Journal of Banking & Finance*, 45, 182-198.

Borio, Claudio and Piti Disyatat (2011): ”Global imbalances and the financial crisis: Link or no link?,” *BIS Working Papers 346, Bank for International Settlements*.

Canzoneri, Matthew, Robert Cumby, Behzad Diba, and David Lopez?Salido (2008): ”Monetary Aggregates and Liquidity in a Neo-Wicksellian Framework,” *Journal of Money, Credit and Banking*, 40(8), 1667-1698.

Christiano, Lawrence J., and Martin Eichenbaum (1992): ”Liquidity Effects and the Monetary Transmission Mechanism,” *The American Economic Review*, 82(2), 346-353.

Clower, Robert (1967): ”A reconsideration of the microfoundations of monetary theory,” *Economic Inquiry*, 6(1), 1-8.

Cooley, Thomas F., and Gary D. Hansen (1991): ”The welfare costs of moderate inflations,” *Journal of Money, Credit and Banking*, Vol. 23, no. 3 : 483-503.

Cooley, Thomas F., and Gary D. Hansen (1989): ”The inflation tax in a real business cycle model,” *The American Economic Review*, 733-748.

Committee on Payment and Settlement Systems (2003): ”The role of central bank money in payment systems,” *Bank for International Settlements*, August 2003.

Cúrdia, Vasco, and Michael Woodford (2016): ”Credit frictions and optimal monetary policy,” *Journal of Monetary Economics*, 84, 30-65.

Del Negro, Marco, Gauti Eggertsson, Andrea Ferrero, and Nobuhiro Kiyotaki (2018): ”The Great Escape? A Quantitative Evaluation of the Fed’s Liquidity Facilities,” *American Economic Review*, 107(3), 824-57

Deutsche Bundesbank (2017): ”The role of banks, non-banks and the central bank in the money creation process,” *Monthly Report April 2017 13*

Disyatat, Piti (2008): ”Monetary policy implementation: Misconceptions and their consequences,” *Bank for International Settlements Working Paper Series*, no. 269.

Disyatat, Piti (2011): ”The bank lending channel revisited,” *Journal of money, Credit and Banking*, 43(4), 711-734.

Durlauf, Steven N., and Lawrence Blume (Eds.) (2008): *The new Palgrave dictionary of economics*, Basingstoke: Palgrave Macmillan.

Galí, Jordi (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press.

Goodfriend, Marvin, and Bennett T. McCallum (2007): "Banking and interest rates in monetary policy analysis: A quantitative exploration," *Journal of Monetary Economics*, 54(5), 1480-1507.

Ireland, Peter N (2004): "Money's Role in the Monetary Business Cycle," *Journal of Money, Credit, and Banking*, 36(6).

Jakab, Zoltan, and Michael Kumhof (2015): "Banks are not intermediaries of loanable funds and why this matters," *Working Paper No. 529. Bank of England*

Kiyotaki, Nobuhiro, and John Moore (2002): "Evil is the root of all money," *The American Economic Review*, 92(2), 62-66.

Kiyotaki, Nobuhiro, and John Moore (2019): "Liquidity, business cycles, and monetary policy," *Journal of Political Economy*, 127, no. 6, 2926-2966.

Fiore, Fiorella De, and Oreste Tristani (2013): "Optimal monetary policy in a model of the credit channel," *The Economic Journal*, 123(571), 906-931.

Lucas, Robert E. (1980): "Equilibrium in a pure currency economy," *Economic inquiry*, 18(2), 203-220.

Lucas, Robert E. (1981): *Models of Business Cycles*, Wiley-Blackwell.

Lucas Jr, Robert E., and Nancy Stokey (1987): "Money and Interest in a Cash-in-Advance Economy," *Econometrica: Journal of the Econometric Society*, 491-513.

Goodfriend, Marvin, and Bennett T. McCallum (2007): "Banking and Interest Rates in Monetary Policy Analysis," *Journal of Monetary Economics*, 54(5), 1480-1507.

McCallum, Bennett T (2001): "Monetary policy analysis in models without money," *Federal Reserve Bank of St. Louis Review*, (Jul), 145-164.

McLeay, Michael, Amar Radia, and Ryland Thomas (2014): "Money in the modern economy: an introduction," *Quarterly Bulletin 2014 Q1*, Bank of England

Ravenna, Federico, and Carl E. Walsh (2006): "Optimal monetary policy with the cost channel," *Journal of Monetary Economics*, 53(2), 199-216.

Ryan-Collins, Josh, Tony Greenham, Richard Werner, and Andrew Jackson (2011): *Where does money come from? A guide to the UK monetary and banking system*, New Economics Foundation.

Stracca, Livio (2007): "Should we take inside money seriously?," *Working Paper Series 0841*, European Central Bank.

Werner, R. A. (2014a): "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*, 36, 71-77.

Werner, R. A. (2014b): "Can banks individually create money out of nothing? The theories and the empirical evidence," *International Review of Financial Analysis*, 36, 1-19.

Woodford, M. (1998): "Doing without money: controlling inflation in a post-monetary world," *Review of Economic Dynamics*, 1(1), 173-219.

Woodford, M. (2003): *Interest and prices: foundations of a theory of monetary policy*, Princeton university press.

Woodford, M. (2008): "How important is money in the conduct of monetary policy?," *Journal of Money, credit and Banking*, 40(8), 1561-1598.

1.8 Appendix

1.8.1 Specification with inputs paid in advance of sales

This section discusses the alternative specification where firms need to pay working capital (here, wages) in advance of sales as in Christiano and Eichenbaum (1992). In this setup the constraints faced by the household slightly change since labor income is received at the beginning of the period: the liquidity-in-advance constraint (1.1) and the equation showing its outstanding wealth at the end of the period (1.2) are now

$$\mu_t : P_tC_t \leq (1 + i_{t-1})D_{t-1} + W_tL_t - D_t^* \tag{1.7}$$

$$\lambda_t : D_t = Div_t + \big((1 + i_{t-1})D_{t-1} + W_tL_t - D_t^* - P_tC_t\big) + D_t^* \tag{1.8}$$

Note that only the expression of the optimality condition with respect to labor changes (i.e. $U_{l,t} - W_t(\mu_t + \lambda_t) = 0$). However, since $\mu_t = 0$, households' behaviour is eventually described by the same equations as in the case presented in section 1.4.3. Households do face, however, liquidity-in-advance constraints (even though slack), hence they purchase goods in exchange of bank deposits.

Firms use bank loans to finance production in advance of sales. Specifically, they exchange their IOUs ("loans") with banks' IOUs ("deposits") which they use to pay workers. Until the end of the period they sell all goods in exchange of bank deposits. Thus, their decision problem *in terms of financial wealth at the beginning of next period* writes[23]:

$$\max_{P_t(i),Y_t(i),L_t(i)} (1 + i_t)P_tY_t - (1 + i_t)W_tL_t \qquad \text{subject to:}$$

$$Y_t(i) = A_tL_t(i)^{1-\alpha} \tag{1.9}$$

$$Y_t(i) = \Big(\frac{P_t(i)}{P_t}\Big)^{-\epsilon} Y_t$$

So, again, their behaviour is described by the same equations as in the cashless version despite the liquidity-in-advance constraints they face.

[23]We follow the approach in the "cost-channel literature" and we assume that firms take one period loans. Note that if we assumed instead that they could repay their loan at the end of the period after receiving proceeds on sales (namely, the case of an intra-temporal loan), the same isomorphy would emerge in the case with inside-money.

Finally, it is important to recall that in contrast to the current inside-money setup, in models with outside-money, the "working-capital in advance" constraint generates an additional "cost channel" of monetary policy transmission (e.g. Ravenna and Walsh (2006), De Fiore and Tristani (2013)) with respect to the basic (credit-frictionless) New-Keynesian model. As a result, the version with "working-capital in advance" is not isomorphic anymore with the latter. This is because firms receive the proceeds on sales in terms of *non-remunerated central bank liabilities*, namely they do not receive any intertemporal (”overnight”) interest on P_tY_t (as it is the case in the setup with 'inside-money', equation (1.9)).

1.8.2 Equations of the model with capital investment

$$\beta(1+i_t)E_t\Big\{\frac{U_{c,t+1}}{U_{c,t}}\frac{P_t}{P_{t+1}}\Big\} = 1$$
$$-\frac{U_{l,t}}{U_{c,t}} = \frac{W_t}{P_t}$$
$$Q_tK_{t+1} - N_{t+1} = D_t$$
$$N_{t+1} = \gamma\Big(Q_tK_t - i_{t-1}\big(Q_{t-1}K_t - N_t\big) - \delta K_tQ_t\Big) + W_t^e + \big(P_tY_t - W_tL_t\big)$$
$$W_t^e = W_t$$
$$C_t^e = (1-\gamma)\Big(Q_tK_t - i_{t-1}\big(Q_{t-1}K_t - N_t\big) - \delta K_tQ_t\Big)$$
$$Y_t = C_t + C_t^e + \mathcal{I}_t$$
$$(1+i_t)Q_t = E_t\left\{\frac{\alpha P_{t+1}Y_{t+1}}{\mathcal{M}K_{t+1}} + Q_{t+1}(1-\delta)\right\}$$
$$K_{t+1} = (1-\delta)K_t + \mathcal{I}_t$$
$$(1-\alpha)\frac{Y_t}{L_t} = \mathcal{M}\frac{W_t}{P_t}$$
$$Y_t = A_tK_t^{\alpha}L_t^{1-\alpha},\ A_t \text{ exogenous}$$
$$Q_t = P_t$$
$$(1+i_t) = \Big(\frac{P_{t+1}}{P_t}\Big)^{\phi}, \phi > 1$$

Chapter 2

MONETARY POLICY WITH FINANCIALLY-CONSTRAINED AND UNCONSTRAINED FIRMS

2.1 Abstract

As monetary policy affects financially-constrained and unconstrained firms to different degrees, its overall impact and design ought to vary with the share of constrained firms in the economy. But it is not clear how. The theoretical literature on the transmission channels and optimal design of monetary policy is ill-equipped to address this question, because it relies on models that feature either constrained firms or unconstrained firms —but not both. In this paper, I enrich the basic New Keynesian model by allowing for both types of firms, and use it to revisit the literature. My model yields a number of novel insights. (i) The interactions of the two types of firms on input and output markets activate a new transmission channel (the "spillover channel"). (ii) Monetary policy affects constrained firms via input prices in the opposite direction of the standard balance-sheet channel (a new "input-price channel"). (iii) Aggregate output does not necessarily respond more strongly to monetary policy when the share of constrained firms is higher (contrary to the financial accelerator intuition), and (iv) "price puzzles" may emerge in equilibrium. (v) Because of the spillover channel, the optimal design of monetary policy does not necessarily change with the share of constrained firms. In the second part of the analysis, I use UK firm-level data to validate the predictions of the model. In the end, I show how the model can be used to discuss the effects of monetary policy in the current low-interest rate environment. Specifically, I show that financially-constrained firms are particularly hurt, whereas unconstrained

firms may benefit when the zero-lower bound on the policy rate binds[1].

Keywords: New Keynesian model, financially-constrained firms, firm heterogeneity, monetary policy transmission, optimal monetary policy

JEL Class.: E2 – E3 – E4.

2.2 Introduction

As monetary policy affects financially-constrained and unconstrained firms to different degrees (figure 2.1), its overall impact ought to vary with the share of constrained firms in the economy. But it is not clear how. The theoretical literature on the transmission channels and design of monetary policy is ill-equipped to address this question because it relies on models that feature either constrained firms or unconstrained firms —but not both. In this paper, I enrich the basic New Keynesian model by allowing for both types of firms, and use it the revisit the literature.

The proposed model is a heterogenous-firm version of the basic NK framework where a share of firms face collateral-constraints, while the others do not face any financing constraints. All firms finance physical and working capital with equity ("net worth") and collateralized debt. Physical capital in fixed aggregate supply ("real estate") serves both as production input and collateral[2]. Firms are heterogenous on one dimension: some of them have high net worth and use it to finance production, while others have low net worth and thus need to rely on (collateralized) debt. The latter end up credit-constrained in equilibrium. I hereafter call these firms for short "constrained". They can be interpreted, for instance, as young firms which lack the net worth and/or performance records required for easy access to credit markets. The model equals the basic NK setup with nominal rigidities à la Rotemberg (1982) on all other dimensions.

[1] I am grateful for support and guidance to my PhD advisors Jordi Galí and Alberto Martin, and for useful comments to Konrad Adler, Frederic Boissay, Andrea Caggese, Davide Debortoli, Egemen Eren, Luca Fornaro, Priit Jenas, Nobuhiro Kiyotaki, Narayana Kocherlakota, Giovanni Lombardo, Rigas Oikonomou, Louis Phaneuf, Stephanie Schmitt-Grohe, Michael Woodford, Raf Wouters, Egon Zakrajsek, Raluca Vernic, Mohammed At Lahcen (discussant SMYE 2018), Irina Marilena Ban (discussant INFER 2018), Pau Belda i Tortosa (discussant BGSE Jamboree 2019), and participants to my presentations at CREi Macro Lunch Seminar, IRES Macro Lunch seminar, Richmond Fed, Kansas City Fed, Bank of Canada, Deutsche Bundesbank, Banque Centrale du Luxembourg, 15th CIREQ PhD conference, 2019 EEA Annual Congress. The empirical analysis is joint work with Ryan Banerjee as part of the BIS PhD Fellowship program. I thank Paolo Surico for sharing the series of monetary policy shocks used in Cloyne et al. (2018).

[2] Real-estate collateral is the working hypothesis of empirical studies on the macroeconomic effects of firms' collateral constraints (Fort et al. (2013), Adelino et al. (2015) for the US, Kleiner (2015), Bahaj et al. (2018) for the UK, Gan (2007) and Lian and Ma (2018) for Japan, and Banerjee and Blickle (2016) and Schmalz et al. (2017) for European countries).

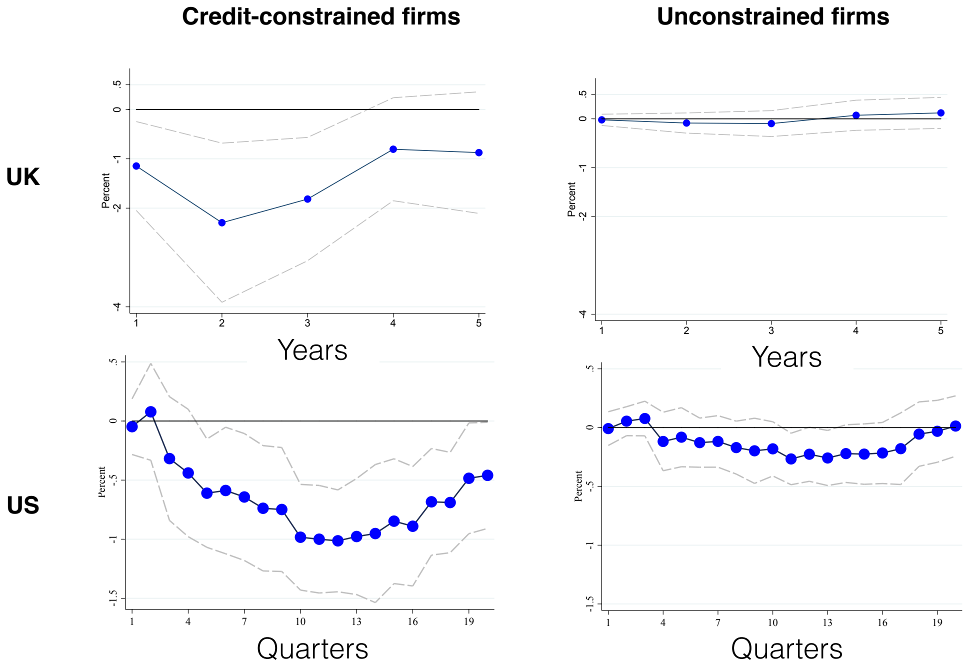


Figure 2.1: Estimated heterogenous responses of investment to a transitory monetary tightening: financially-constrained *versus* unconstrained firms in the UK (top panels) and the US (bottom panels)
Source: Cloyne et al. (2018)

Firms' collateral constraints activate a number of new monetary policy transmission mechanisms. In the basic version without financial frictions, production by all firms is affected via the aggregate demand channel: as individual demand schedules decline in response to a rise in the interest rate, firms unequivocally reduce both prices and output levels. In the presence of credit frictions however, a monetary tightening affects production by constrained firms via additional mechanisms in opposite directions. On the one hand, it affects production by constrained firms *negatively* by pushing downwards collateral values and upwards the real value of nominal debt (the standard balance-sheet and debt deflation channels). On the other hand, it affects it *positively* because it reduces prices of inputs financed against collateral (*an input price channel* not considered so far in the literature). Effects on constrained firms further spill over to unconstrained ones via input and output markets, either enhancing or dampening the direct effect of monetary policy on the latter. These "spillovers" constitute a seconds new channel of monetary policy transmission which has been overlooked in the literature.

Other three interesting conclusions emerge from the analysis of the effect of firms' credit frictions on aggregate transmission. First, contrary to conventional wisdom, given the opposing nature of mechanisms affecting output of constrained firms and the existence of spillovers, credit frictions may both amplify or dampen the response of aggregate output to monetary policy. A strong balance-sheet channel favors amplification, whereas a strong input-price channel favors dampening. Second, because of spillovers (and hence, general equilibrium effects), observing constrained firms responding more on average (as for instance in the UK or the US) does not necessarily imply that credit frictions significantly amplify the macroeconomic response. This is because, while constrained firms are severely affected and respond more than a firm in a world without financial frictions, the unconstrained ones relatively benefit, and their activity reduces less than in the absence of credit frictions. Third, since monetary policy affects not only aggregate demand but also the supply of constrained firms, it may have an unexpected impact on prices. Specifically, a contractionary monetary policy may depress supply strongly enough together with demand to induce a rise in the price level in equilibrium (instead of a decline as in the basic model).

The changes in monetary transmission have also implications for the optimal design of monetary policy. Given its empirical relevance, I focus exclusively on the case of a dominant balance-sheet channel. I find that spillovers play an important role and hence the optimal policy response is very different from the weighted mean of the two polar limiting cases (with only unconstrained, or with only constrained firms). Furthermore, according to the analysis, the strength of the balance-sheet channel affects decisively the optimal response to non-financial shocks. Specifically, the stronger this channel, the weaker the reduction of the policy rate under optimal policy in response to negative demand and positive technology shocks. This is because in the presence of constrained firms and of a dominant balance-sheet channel, the decline in the policy rate is associated not only to an upward shift in aggregate demand, but also in aggregate supply. Thus, relatively to a credit-frictionless environment, in response to a negative demand shock or to a positive technology shock, a monetary loosening induces additional deflationary pressures via its supply side effects. Thus, when the balance-sheet channel is strong, contrary to the credit-frictionless benchmark, a decline in the policy rate that allows to close the output gap, may be associated to inefficient deflationary pressures in equilibrium.

In response to an adverse financial shock, modeled as an exogenous reduction in the pledgeability of capital as collateral, I find that the policy rate declines so as to prop up collateral asset prices under optimal monetary policy. This decline is aimed at (partially) counteracting the effect of the shock. A similar result is obtained by Andres et al. (2013) based on the setup in Iacoviello (2005) which features a strong balance-sheet channel in the context of collateral constraints, as well as by De Fiore and Tristani (2013) in the context of costly-state verification for a shock to constrained firms's net worth. The latter analyses however do not take into account the transmission spillovers between the two types of firms and the effect of steady-state distortions. Thus, relatively to them, in the current heterogenous firm setup we can also see how the decline in the policy rate further induces unconstrained firms to expand production, and to push output above its inefficient steady-state level.

In the second part of the paper, I use firm-level data to test the theoretical predictions of the model. For this purpose, I follow Cloyne et al. (2018) and I use firm balance-sheet data from the WorldScope database for the UK. Constrained firms are defined as young firms which do not distribute dividends. I estimate the responses of constrained versus unconstrained firms to monetary policy using an instrumental variable version of the local projection method developed by Jorda et al. (2019), with the interest rate instrumented by high frequency monetary policy shocks from Gerko and Rey (2015). Theoretical predictions of the model regarding monetary policy transmission are consistent with the data. In particular, the model predicts that all else equal, a lower pledgeability of capital or a lower liquidity ratio are associated to a stronger response of constrained firms to monetary policy in the conventional direction, which is corroborated by the data. Moreover, at least for a subset of constrained firms, a monetary tightening appears to steer their activity in an unconventional positive direction. This happens for those with only short-term debt (used as a proxy for working capital credit) and a low fraction of tangible capital.

In the third part of the paper, I discuss how the model can be used for policy analysis. Given its relevance at the moment, I focus on the particular case of a low interest rate environment. I start by interpreting the estimated impact of monetary policy on constrained versus unconstrained firms in the UK through the lenses of the model. Estimations show that constrained firms reduce their activity stronger than unconstrained ones in response to a monetary tightening. These results are consistent with a dominant balance-sheet channel. Specifically, the increase in the policy rate reduces collateral asset values, and this makes constrained firms cut strongly production. The strong adverse effects on constrained firms have positive spillovers for unconstrained ones, and these positive spillovers partially counteract the negative effects of the monetary tightening on unconstrained firms. Hence, the latter respond only mildly in equilibrium.

I then show how in economies with strong dominant balance-sheet channels such as the UK (and also the US and most likely the Euro Area), the ZLB on the policy rate hurts constrained firms, and may benefit unconstrained ones. Specifically, in the absence of the ZLB, constrained firms are more negatively affected than unconstrained ones in response to a monetary tightening, but more positively affected in response to a loosening. Thus, in the absence of the ZLB, the relatively gains and losses for these firms compensate over the business cycle. Since the

ZLB limits the more positive effects of monetary policy when a decline in the policy rate is warranted, but the more adverse effects remain unchanged, on average over the business cycle, constrained firms end up particularly hurt. Furthermore, because of spillover effects, unconstrained firms benefit. Specifically, since constrained firms are affected less positively by a cut in interest rates, their negative spillovers to unconstrained ones via input and output markets are also lower. As a result, monetary policy has stronger net positive effect on unconstrained firms in times when a cut in the policy rate is needed. Thus, despite the stronger decline in aggregate activity when the ZLB binds, unconstrained firms may not be affected by the latter in equilibrium, or they may even end up producing more.

Hereafter, section 2 reviews related literature, section 3 describes the model, section 4 analyses monetary policy transmission, and section 5 focuses on optimal design. The paper concludes by discussing future extensions.

2.3 Relation to the literature

The paper is related to three main strands of literature. The first one is the theoretical literature on the transmission and optimal design of monetary policy in the absence of financial frictions (Galí (2015), Woodford (2003)). In this world, monetary policy transmits to the economy by shifting its aggregate demand channel. It can thus always insulate it from the (welfare loss) effects of demand shocks, which is equivalent to varying the policy rate so as to achieve zero inflation in equilibrium. By contrast, in response to technology shocks, this literature finds that monetary policy should optimally target a composite measure of price and wage inflation. This is because the efficient real wage varies, and this variation is approached via a (costly) adjustment in both sticky prices and wages. Finally, financial shocks play no role in this frameworks.

The second strand of literature is the one studying the implications of firms' financial frictions for monetary policy[3]. With two exceptions (Gilchrist et al. (2017) and Ottonello and Winberry (2018)), all models used so far feature only financially-constrained firms (Iacoviello (2005), Carlstrom et al. (2010), Andres et al. (2013) for collateral constraints, De Fiore and Tristani (2013), Faia and Monacelli (2007), Hansen (2018) and Fendoglu (2014) for other types of financial frictions). The two exceptions are models built to rationalize particular stylized facts identified in US data, namely (i) why financially-constrained firms increased prices during the recent financial crisis, while the unconstrained decreased them, and (ii) why low-leverage firms respond more to monetary policy. Relatively to

[3]An extensive strand of theoretical literature starting with Kiyotaki and Moore (1997) studies how such financial frictions may alter macroeconomic dynamics within flexible-price ('real') frameworks. These models however do not consider monetary policy.

these two models which are tailored to deal with very specific questions, the model in this paper is more general. Thus, *inter-alia*, it can be used also to rationalize the two particular patterns. Specifically, the first pattern emerges as well in the current monopolistic competition setup (in the absence of customer-based markets as assumed by the former paper), whereas the second arises to the extent that low-leveraged firms are financially-constrained firms with low collateral.

Furthermore, papers in this literature have either considered working capital credit or investment credit, whereas the current analysis considers both[4]. The inclusion in the model of both types of credit enables to reconcile seemingly contradictory findings of previous studies on the implications of financial frictions for the macroeconomic effects of monetary policy. For instance, collateral constraints strongly amplify the effect of monetary policy on output in Iacoviello (2005) and Andres et al. (2013) (which considered the case of investment credit), whereas they hardly have any effect in Carlstrom et al. (2010) or De Fiore and Tristani (2013) (which considered the case of working-capital credit). Through the lenses of the current setup, the first result is explained by a strong balance-sheet channel (embedded in the first two models via a financial accelerator à la Kiyotaki and Moore (1997)), whereas the second by a strong input-price channel counteracting the balance-sheet one. In the current analysis both instances are possible in equilibrium depending on structural parameters.

Another distinguished feature of the current setup is that production and price decisions are taken jointly at the monopolistic firm level (as in the basic NK model). And this irrespectively of whether the firm is financially-constrained or not. This allows to study how monetary policy affects jointly firm-level output and prices. By contrast, with the exception of Gilchrist et al. (2017), in all other models production decisions are taken by perfectly competitive wholesalers, whereas pricing decisions are taken separately by monopolistic retailers which use the output of wholesalers as input. The setup in Gilchrist et al. (2017) is different than the current one because it looks at the particular case of customer-based markets instead of a standard monopolistic competition environment.

Finally, the paper contributes to the empirical literature on the effect of monetary policy on financially constrained versus unconstrained firms. Previous papers distinguished between the (average) effect of monetary policy on the two groups, using different proxies for financial constraints (e.g. young and not distributing dividends in Cloyne et al. (2018), high leverage in Ottonello and Winberry (2018), small in Gertler and Gilchrist (1994)). So far, I used the model as a guide to study

[4]As argued by Jermann and Quadrini (2012), while the relevance of investment credit is obvious, working capital credit helps rationalizing the strong link between collateral and employment observed in the US data. As shown later in the analysis, it also allows to explain the estimated effect of monetary policy on constrained firms in the UK (in particular, the role of capital tangibility).

in more detail the transmission of monetary policy to a subset of UK (public) firms which are likely to be credit-constrained, by extending the analysis in Cloyne et al. (2018). Going forward, I aim to further contribute to this literature by bringing supportive evidence for the new transmission channels of monetary policy identified in the theoretical analysis, namely for the "input price channel" and the "spillover channel".

2.4 Model

The analytical framework takes as a benchmark the Rotemberg version of the basic NK setup with working-capital paid in advance of sales[5]. This basic setup features three types of agents: a continuum of identical households which consume, work and save, differentiated monopolistic firms which produce, and a monetary authority which sets the one-period nominal interest rate. I add two types of financial frictions to this setup such that a set of firms end up credit-constrained in equilibrium. Without these frictions, the setup equals the basic model on all other dimensions[6].

I assume that there are two types of firms in the economy: firms which do not face any financing frictions ("unconstrained"), and firms which face two types of financial frictions ("constrained"). The first financial friction is a limit on their net worth. Specifically, these firms can finance with equity only a fraction of their desired physical capital (due to high issuance costs related to moral hazard for instance)[7]. For tractability, as in Gilchrist et al. (2017), constrained firms issue each period equity claims and do not retain profits[8]. The second financial friction is the requirement to secure nominal debt against collateral[9]. Because

[5]The Calvo-version of the basic NK model (which is equal to a first order approximation to the Rotemberg version) is extensively described in Gali (2015) and Woodford (2003). In this baseline setup, working capital constraints do not affect equilibrium (Manea (2019)).

[6]Papers so far depart from the basic NK setup, not only by adding credit frictions, but also by assuming households and firms have distinct preferences. In Iacoviello (2005) and Andres et al. (2013) 'patient' households work, consume and save in equilibrium, whereas 'impatient' entrepreneurs produce, consume and borrow. In Carlstrom et al. (2010), households are risk-averse, whereas entrepreneurs are risk-neutral. Moreover, households supply two types of labor (one on which firms face credit collateral constraints).

[7]'Equity' should thus be understood more generally in the model as firm's *net worth*, namely as both fresh equity injections ('external equity') and retained earnings (internal equity).

[8]This allows to avoid technical difficulties due to heterogenous net worth (hence, output) of constrained firms.

[9]To microfound the latter friction, we may think of firms as being run by a special category of workers, 'managers', who can refuse to repay debt so as to maximize firms' shareholders' revenues. Creditors are thus only willing to lend against physical capital that can be seized and liquidated in case of repudiation.

of these two frictions, the second type of firms face limits on how much they can produce. We can think of them as being relatively young, and thus without sufficient retained earnings so as to finance internally their operations, and without a well-established credit record so as to get external financing in an unconstrained manner. All constrained firms are identical and produce an identical amount. The same is true for unconstrained firms. I take the size of each set as exogenous. I note the size of the constrained group by ϕ, and of the unconstrained one by $1-\phi$. Firm entry and exit flows are such that the distribution of the two types of firms is constant over time[10].

There are five markets in the model: goods, labor, capital, debt and equity. We may think of debt contracts as being intermediated via the banking sector. Thus, monetary authority's instrument is the interest rate on one-period nominal debt.

2.4.1 Households

The economy is populated by a continuum of identical infinitely-lived households of measure one. At each date, a representative household decides how much to consume (C_t), to work (L_t), and to invest in one-period nominal debt $(\mathcal{D}_t)$ and equity shares $(\{\mathcal{E}_t(i)\})$ issued by firms, in order to maximise expected lifetime utility

$$E_0\Big\{\sum_{t=0}^{\infty}\beta^t Z_t\Big(\frac{C_t^{1-\sigma}-1}{1-\sigma}-\frac{L_t^{1+\varphi}}{1+\varphi}\Big)\Big\}$$

subject to the sequence of budget constraints.

$$W_tL_t+\big(1+i_{t-1}\big)\mathcal{D}_{t-1}+\int_0^1\mathcal{E}_{t-1}(i)\mathcal{R}^e_{t-1}(i)di\geq P_tC_t+\mathcal{D}_t+ \tag{2.1}$$

$$+\int_0^1\mathcal{E}_t(i)Q^e_t(i)di+T_t \tag{2.2}$$

and the solvency ('transversality') conditions

$$\lim_{T\to\infty}E_0\Big\{\beta^T\frac{U_{c,T}}{U_{c,t}}\frac{\mathcal{D}_T}{P_T}\Big\}\geq 0,\qquad \lim_{T\to\infty}E_0\Big\{\beta^T\frac{U_{c,T}}{U_{c,t}}\frac{\mathcal{E}_T(i)}{P_T}\Big\}\geq 0, \tag{2.3}$$

with $C_t\equiv\Big[\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon}}di\Big]^{\frac{\varepsilon}{\varepsilon-1}}$ a standard Dixit-Stiglitz consumption index of differentiated goods with ε a measure of substitutability among them and its unit

[10]Namely, (i) at each date a mass of newborn firms enter the constrained set, while a mass of equal size simultaneously unexpectedly exit it to enter the (well-established) unconstrained set; (ii) simultaneously, a mass of equal size with the firms entering the unconstrained set unexpectedly exits the economy.

price denoted by $P_t \equiv \int_0^1 \Big(P_t(i)^{1-\varepsilon} di\Big)^{\frac{1}{1-\varepsilon}}$, W_t the nominal wage, i_t the one-period interest rate on nominal debt, $Q_t^e(i)$ the price of an equity claim in firm i, $\mathcal{E}_t(i)$ the number of equity claims in firm i, T_t (lump-sum) monopolistic profits distributed by firms, and Z_t an exogenous demand preference shifter described by $log(Z_t) = \rho_z log(Z_{t-1}) + \varepsilon_t^z,\ \varepsilon_t^z \sim \mathcal{N}(0, \sigma_z)$[11]. Let $\Lambda_{t,t+1} \equiv \beta\Big(\frac{Z_{t+1}}{Z_t}\Big)\Big(\frac{C_{t+1}}{C_t}\Big)^{-\sigma}$ denote the stochastic discount factor for one-period ahead (real) payoffs. Representative household's behavior is described by

$$C_t^{\sigma} L_t^{\varphi} = \frac{W_t}{P_t} \quad \forall t, \tag{2.4}$$

$$E_t\Big\{\Lambda_{t,t+1}\Pi_{t+1}^{-1}\Big\}(1+i_t) = 1 \quad \forall t, \tag{2.5}$$

$$E_t\Big\{\Lambda_{t,t+1}\Pi_{t+1}^{-1}\mathcal{R}_t^e(i)\Big\} = Q_t^s(i) \ \ \forall i \ \ \forall t, \tag{2.6}$$

alongside the budget constraints (2.1), and the transversality condition (2.3).

2.4.2 Firms

The model economy is populated by a continuum of firms in monopolistic competition which produce differentiated goods indexed by $i \in [0, 1]$. At each date, firms have access to an identical constant returns to scale Cobb-Douglas technology

$$Y_t(i) = A_t K_t^{\alpha}(i) L_t^{1-\alpha}(i),$$

where Y stands for output, K for capital, and $log(A_t) = \rho_a log(A_{t-1}) + \varepsilon_t^a$, $\varepsilon_t^a \sim \mathcal{N}(0, \sigma_a)$ is a common stochastic productivity process. Capital is in fixed aggregate supply.

Firm i enters period t with predetermined capital $K_t(i)$ chosen at the end of $t-1$, and purchased at $t-1$ on a perfectly competitive market at price Q_{t-1}^k. Firms refinance each period their entire capital. The fraction $\theta(i)$ of $K_t(i)$ was financed by one-period equity claims at time $t-1$. Firm i issued at that time $\mathcal{E}_t(i)$ claims equal to the number of capital units financed by equity, and priced each claim at the price of a capital unit, namely $Q_{t-1}^e = Q_{t-1}^k$[12], while it financed the rest $\big(1-\theta(i)\big)Q_{t-1}^k K_t(i)$ by nominal debt $\mathcal{D}_{t-1}(i)$ at interest rate i_{t-1}[13].

Firms need to pay the wage-bill in advance of sales ("cash-flow mismatch"). Subsequently, they need to finance each period not only physical capital, but also

[11] As households are the patent owners of firm technology, they earn monopolistic profits.

[12] The same approached is followed in Gertler and Karadi (2011).

[13] As in Iacoviello (2005), nominal debt (as opposed to inflation-indexed one) is justified by the observation that in low inflation countries almost all debt contracts are in nominal terms.

working-capital in advance of sales[14]. As in Jermann and Quadrini (2012) or Carlstrom et al. (2010)), they do so by issuing intratemporal (interest-free) debt. Firms can only issue debt against collateral. Firm's i total debt cannot exceed a pledgeable fraction ν_t of its capital holdings (that can be seized in case of repudiation) at the end of the period (when debt is repaid):

$$W_t L_t(i) + (1 + i_{t-1})\mathcal{D}_{t-1}(i) \leq \nu_t Q_t^k K_t(i), \quad log(\nu_t) = \rho_\nu log(\nu_{t-1}) + \varepsilon_t^{\nu 15} \tag{2.7}$$

The sequence of events is as follows: at the beginning of t, subject to the credit collateral constraint (2.7), firm i decides how much to produce $Y_t(i)$, hires workers $L_t(i)$ and sets the price $P_t(i)$ in the presence of Rotemberg (1982)-style adjustment costs $\zeta_t(\cdot) \equiv \frac{\xi}{2} Y_t \Big(\frac{P_t(i)}{P_{t-1}(i)} - 1 \Big)^2 \geq 0$[16]. Once it receives its sales income at the end of the period, it resells capital, repays debt and equity returns, and redistributes monopolistic profits to households. After honoring all liabilities related to current production, it chooses capital for next period $K_{t+1}(i)$, and the production cycle starts over again.

Formally, at date t, firm i takes as given its capital $K_t(i)$, its equity financing constraint $\theta(i)$, the wage W_t, the price level P_t, aggregate demand Y_t, and chooses $\{P_t(i), Y_t(i), L_t(i), K_{t+1}(i), \mathcal{D}_t(i)\}_{t \geq 0}$ to maximize its expected intertemporal profits:

$$E_0 \Bigg\{ \sum_{t=0}^{\infty} \Lambda_{0,t} \Bigg[\frac{P_t(i)}{P_t} Y_t(i) - \frac{W_t}{P_t} L_t(i) - \frac{\xi}{2} Y_t \Big(\frac{P_t(i)}{P_{t-1}(i)} - 1 \Big)^2 + \frac{Q_t^k}{P_t} K_t(i) + \frac{\mathcal{D}_t(i)}{P_t} - \Big(1 + i_{t-1}\Big) \Pi_t^{-1} \frac{\mathcal{D}_{t-1}(i)}{P_{t-1}} - R_{t-1}^e \frac{\mathcal{E}_{t-1}(i)}{P_t} - \frac{Q_t^k}{P_t} K_{t+1}(i) \Bigg] \Bigg\},$$

subject to the sequence of collateral constraints:

$$\lambda_t^1(i) : \nu_t \frac{Q_t^k}{P_t} K_t(i) - (1 + i_{t-1}) \frac{\mathcal{D}_{t-1}(i)}{P_{t-1}} \Pi_t^{-1} \geq \frac{W_t}{P_t} L_t(i) \;\; \forall t, \tag{2.8}$$

demand constraints:

$$\lambda_t^2(i) : \Bigg(\frac{P_t(i)}{P_t} \Bigg)^{-\varepsilon} Y_t - Y_t(i) = 0 \;\; \forall t, \tag{2.9}$$

[14] As working capital loan is intra-temporal (as in Jermann and Quadrini (2012) or Carlstrom et al. (2010)), the effects of their interest rate on the credit constraint tightness are ignored.

[15] If firms repudiate their debt obligations, banks can reposes their collateral assets only by paying a proportional transaction cost (e.g. Quadrini (2011)), so $\nu_t < 1$.

[16] I chose Rotemberg (1982)-style nominal price rigidities because they allow price and quantity decisions to be jointly taken by standard monopolistic firms in the presence of collateral constraints (Jermann and Quadrini (2012)).

technology constraints:

$$\lambda_t^3(i) : Y_t(i) - A_t K_t^{\alpha}(i) L_t^{1-\alpha}(i) = 0 \ \ \forall t, \tag{2.10}$$

and budget constraints[17]

$$\begin{aligned}&\Bigg[\frac{P_t(i)}{P_t}Y_t(i) - \frac{W_t}{P_t}L_t(i) - \frac{\xi}{2}Y_t\Big(\frac{P_t(i)}{P_{t-1}(i)} - 1\Big)^2 + \frac{Q_t^k}{P_t}K_t(i) - \Big(1+i_{t-1}\Big)\frac{\mathcal{D}_{t-1}(i)}{P_t}\\ &- \frac{R_{t-1}^e(i)\mathcal{E}_{t-1}(i)}{P_t} - \frac{T_t}{P_t}\Bigg] + \frac{\mathcal{D}_t(i)}{P_t} + \frac{Q_t^e(i)\mathcal{E}_t(i)}{P_t} = \frac{Q_t^k}{P_t}K_{t+1}(i) \ \ \forall t.\end{aligned} \tag{2.11}$$

where $\lambda_t^1(i)$, $\lambda_t^2(i)$ and $\lambda_t^3(i)$ are the lagrange multipliers associated to each of the three binding constraints in firm (i)'s maximization problem. Since firm i fully repays gross equity returns and monopolistic profits at the end of each period (the term in brackets is zero), and equity issued at t equals $\theta(i)Q_t^k K_{t+1}(i)$, firm i's budget constraint (2.11) implies:

$$\mathcal{D}_t(i) = (1-\theta(i))Q_t^k K_{t+1}(i) \tag{2.12}$$

We can thus replace the expression of $\mathcal{D}_t(i)$ from (2.12) in firm i's optimization problem, and eliminate budget constraint (2.11) alltogether. The Lagrangian method gives the following optimality conditions for firm i's behaviour:

$$\xi\Pi_t(i)\Big(\Pi_t(i)-1\Big)\frac{Y_t}{Y_t(i)} = \xi E_t\Big\{\Lambda_{t,t+1}\Big[\frac{Y_{t+1}}{Y_t(i)}\Pi_{t+1}(i)\Big(\Pi_{t+1}(i)-1\Big)\Big]\Big\} + \Big(\frac{P_t(i)}{P_t} - \varepsilon\lambda_t^2(i)\Big),$$

$$\frac{P_t(i)}{P_t} - \lambda_t^2(i) - MC_t(i)\Big(1+\lambda_t^1(i)\Big) = 0,$$

$$\begin{aligned}\frac{Q_t^k}{P_t} =& E_t\Bigg\{\Lambda_{t,t+1}\Bigg[\frac{Q_{t+1}^k}{P_{t+1}} + \Big(\frac{P_{t+1}(i)}{P_{t+1}} - \lambda_{t+1}^2(i)\Big)\alpha\frac{Y_{t+1}(i)}{K_{t+1}(i)} +\\ &+ \lambda_{t+1}^1(i)\Big(\nu_{t+1}\frac{Q_{t+1}^k}{P_{t+1}} - (1+i_t)(1-\theta(i))\frac{Q_t^k}{P_t}\Pi_{t+1}^{-1}\Big)\Bigg]\Bigg\},\end{aligned}$$

$$\lambda_t^1(i)\Big(\nu_t\frac{Q_t^k}{P_t}K_t(i) - \Big(1+i_{t-1}\Big)\frac{\mathcal{D}_{t-1}(i)}{P_{t-1}}\Pi_t^{-1} - \frac{W_t}{P_t}L_t(i)\Big) = 0, \ \ \lambda_t^1(i) \geq 0$$

$$\nu_t\frac{Q_t^k}{P_t}K_t(i) - \Big(1+i_{t-1}\Big)\frac{\mathcal{D}_{t-1}(i)}{P_{t-1}}\Pi_t^{-1} - \frac{W_t}{P_t}L_t(i) \geq 0, \ \ \forall t,$$

alongside the sequence of demand (3.17) and technological (2.10) constraints. $MC_t(i) \equiv \frac{1}{1-\alpha}\frac{W_t}{P_t}\frac{L_t(i)}{Y_t(i)}$ denotes firms i's real marginal cost.

[17] According to the budget constraint, capital is financed with new debt, new equity and previous capital returns net of old debt, equity returns and redistributed monopolistic profits.

The model economy is populated by two sets of firms: a set of mass ϕ denoted by Θ^c with $\theta(i)$ low enough for collateral constraints to bind in the vicinity of steady-state, and another set denoted by Θ^u of mass $1-\phi$ with $\theta(i)=1$. Firms within each set behave identically. Variables related to the constrained set are indexed by $'c'$, and the ones to the unconstrained one by $'u'$.

2.4.3 Monetary authority

Unless otherwise stated, the monetary authority sets the nominal interest rate

$$\widehat{i}_t = \phi_\pi \pi_t + \phi_y \widehat{y}_t + \varepsilon_t^m, \tag{2.13}$$

where $\pi_t \equiv log(\Pi_t)$, $\widehat{y}_t \equiv log(Y_t) - log(Y)$ and $\varepsilon_t^m = \rho_m \varepsilon_{t-1}^m + \epsilon_t^m$, $\epsilon_t^m \sim \mathcal{N}(0, \sigma_m)$.

2.4.4 Market clearing

Market clearing is imposed:

- for each good variety i: $Y_t(i) = C_t(i) + \xi_t(i), \ \forall i$ with $\xi_t(i)$ the price adjustment costs in terms of variety i[18];

- on labor market where labor supplied by households must equate labor demanded by firms: $L_t = \int_0^1 L_t(i)di = \phi L_t^c + (1-\phi)L_t^u$;

- on capital market where firms' aggregate demand must equate the (exogenously fixed) aggregate supply: $\bar{K} = \int_0^1 K_{t+1}(i)di = \phi K_{t+1}^c + (1-\phi)K_{t+1}^u$;

- on debt market where demand equals supply subject to collateral requirements, for both inter-temporal debt $\mathcal{D}_t = \int_0^\phi (1-\theta^c) Q_t^k K_{t+1}(i)di = \phi(1-\theta^c) Q_t^k K_{t+1}^c$, and intra-temporal one $\mathcal{D}_t^i = W_t L_t$;

- on equity market where the value of equity claims demanded by households must equate the value issued by constrained and unconstrained firms[19]:

$$\int_0^1 \mathcal{E}_t(i) Q_t^e(i) di = \int_0^\phi Q_t^e(i)\mathcal{E}_t(i)di + \int_\phi^1 Q_t^e(i)\mathcal{E}_t(i)di \tag{2.14}$$

$$= \phi\theta^c Q_t^k K_{t+1}^c + (1-\phi) Q_t^k K_{t+1}^u \tag{2.15}$$

[18]The allocation of adjustment costs among varieties is the same as for consumption.

[19]Households are the ultimate owners of all firms in the economy.

2.5 Monetary policy transmission

I split the theoretical analysis of the transmission of monetary policy in the presence of constrained and unconstrained firms in three parts. In the first part, I investigate how the financial constraints faced by firms shape the mechanisms of transmission of monetary policy relatively to the credit-frictionless benchmark in Galí (2015) or Woodford (2003). In the second part, I analyze how the effect of monetary policy at the firm level depends on structural parameters, focusing on examples that can be tested empirically. In the third stage, I study how the reaction of *macroeconomic* variables is affected by the share of constrained firms, and how results hinge on structural parameters. I base my theoretical analysis on a first order log-linear approximation of the model in the vicinity of the non-stochastic zero-inflation steady-state[20]. Notation is standard: small caps stand for log-levels, $(\widehat{\ })$ for log-deviation from steady-state, while the absence of a time subscript denotes a steady-state value.

2.5.1 Transmission mechanisms

Let's consider first the limiting case of $\phi = 0$. Given nominal rigidities à la Rotemberg, all firms choose the same price and output levels. Equilibrium dynamics may be summarized by the following three equations:

$$\widehat{y}_t = E_t\{\widehat{y}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\}\Big) + (1-\rho_z)z_t \tag{2.16}$$

$$\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\widehat{y}_t - \lambda\frac{1+\varphi}{1-\alpha}a_t \tag{2.17}$$

$$\widehat{i}_t = \phi_\pi \pi_t + \phi_y \widehat{y}_t + \varepsilon_t^m \tag{2.18}$$

For ξ calibrated such that λ equals the value in the Calvo version, the model is isomorphic (up to a first order) to the basic NK setup in Gali (2015, Chapter 3).

A monetary tightening reduces aggregate demand and is transmitted to firms' decisions through associated declines in their individual demand schedules[21]. In the credit-frictionless environment, firms can produce as much as they want subject to their demand and price setting constraints. Figure 2.2 depicts the response of a firm to a monetary tightening in this limiting case. Specifically, it shows how as the negative monetary impulse reduces its demand schedule, namely it shifts

[20] Shocks are small enough for supply to be non-rationed and credit constraints to bind.

[21] Output and inflation dynamics are independent of asset prices in the credit-frictionless limit. For a detailed analysis of monetary policy transmission and optimal design in this case see for instance Gali (2015), Chapters 3 to 5.

it to the left from D_0 (black solid line) to D_1 (navy dotted line), the firm simultaneously readjusts downwards its output (from Q_0 to Q_1) and price (from P_0 to P_1) given the marginal cost schedule (Cmg line) and price adjustment costs[22]. In this environment, the (real) marginal cost schedule of the economy is also affected in equilibrium. However, its endogenous variation is never strong enough so as to overturn at the firm-level the sign of the direct effect of the monetary impulse via the demand channel. This is why, the "input-price channel" does not play a key independent role in the transmission of monetary policy when the economy is populated only by unconstrained firms.

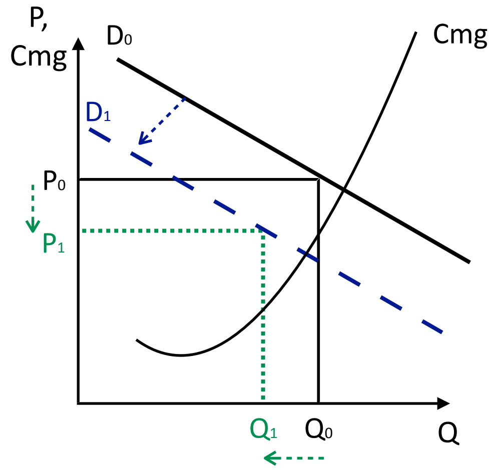


Figure 2.2: Firm response to a monetary tightening in the credit-frictionless limit

In the general case when some of the firms are constrained, namely $\phi \in (0, 1)$, monetary impulses are still directly exerted via the demand/saving channel. This is because the structure of the demand-side of the model economy remains the same as the one in the credit-frictionless limit. A series of indirect effects acquire however new roles in the transmission of monetary policy at the firm level. Moreover, monetary policy transmits to constrained firms through different mechanisms than to unconstrained ones, and spillover effects between the two types firms play a key role in the transmission of monetary policy.

At each date, the output of *a constrained firm* is determined by its collateral constraint (2.30). This constraint implies that given outstanding capital $\widehat{k}_t^c$ and nominal debt (and exogenous disturbances), the production of such a firm depends

[22]Firm's real marginal cost schedule also shifts in equilibrium, but it is held fixed in figure 2.2 for ease of exposition. Price adjustment costs (not modeled) imply that the firm takes into account as well future marginal cost and demand schedules.

on (i) the current price of capital pledged as collateral $\widehat{\varrho}_t^k$, (ii) price inflation π_t, and (iii) the prices of the inputs financed against collateral (here, real wage $\widehat{\omega}_t$):

$$\begin{aligned}\boldsymbol{\widehat{y}_t^c} &= \widehat{k}_t^c - \frac{(1-\alpha)\beta^{-1}(1-\theta^c)}{\nu-\beta^{-1}(1-\theta^c)}\left[\widehat{\varrho}_{t-1}^k + \left(\widehat{i}_{t-1}-\pi_t\right)\right] \\ &+ \frac{(1-\alpha)\nu}{\nu-\beta^{-1}(1-\theta^c)}\left(\widehat{\varrho}_t^k - \nu_t\right) - (1-\alpha)\,\widehat{\omega}_t + a_t\end{aligned}$$

Thus, a transitory monetary impulse transmits to the current production of a constrained firm via its indirect effects on (i) the real price of capital pledged as collateral, (ii) the real value of outstanding (nominal) debt (via inflation), and (iii) the real prices of the inputs financed against collateral:

$$\frac{\partial \widehat{y}_t^c}{\partial \varepsilon_t^m} = \overbrace{\frac{(1-\alpha)\nu}{\nu-\beta^{-1}(1-\theta^c)}\frac{\partial \widehat{\varrho}_t^k}{\partial \varepsilon_t^m}}^{\text{balance-sheet}} + \overbrace{\frac{(1-\alpha)\beta^{-1}(1-\theta^c)}{\nu-\beta^{-1}(1-\theta^c)}\frac{\partial \pi_t}{\partial \varepsilon_t^m}}^{\text{nominal-debt}} \overbrace{-(1-\alpha)\frac{\partial \widehat{\omega}_t}{\partial \varepsilon_t^m}}^{\text{input-price}}$$

The first mechanism is a facet of the standard "balance-sheet channel" which refers to the fact that a rise in the policy rate depresses asset prices, and hence shrinks the value of firm's collateral[23]. The second one is the "nominal debt-channel": since (outstanding) debt is nominal, monetary policy affects its real value via inflation[24]. Both mechanisms have been already discussed in the literature (e.g. the first one in Bernanke and Gertler (1995), both in Iacoviello (2005)). The third one however, working via the real prices of the inputs financed against collateral, to the best of my knowledge, has been overlooked so far. Its key relevance in this context is given by the consideration of working-capital credit secured against collateral. Firm heterogeneity plays also an important role in determining its strength in equilibrium. I call this new channel the *input-price channel*.

Importantly, these three transmission mechanisms affect production by constrained firms in opposite directions. For instance, a transitory monetary tightening simultaneously *lowers* the production possibilities of these firms by pushing downwards the price of their pledgeable collateral via the "balance-sheet" mechanism, and *expands* them by depressing real input prices (via the decline in output, hence, in input demand). Depending on how inflation reacts in equilibrium, "debt-deflation" might further affect them in either direction. Thus, *a priori*, monetary

[23]Bernanke and Gertler (1995) note that 'many observers would agree that the crash of Japanese land in the latter 1980s was the result (at least in part) of monetary tightening and that this collapse in asset values reduced the credit-worthiness of many Japanese corporations, contributing to the ensuing recession.' They also mention that according to Borio et al. (1994) 'a similar pattern of asset price boom and bust leading to real fluctuations occurred during the 1980's in a number of major industrialized countries'.

[24]The nominal debt channel plays an important role in the 'debt-deflation' theory of the (1929-1939) 'Great Depression in US' proposed by Fisher (1933).

policy may steer current output of constrained firms in either direction depending on the relative strength of each mechanism (and hence, on structural parameters).

These three transmission mechanisms are further relevant for the pricing decisions of constrained firms. This is because their prices depend, alongside aggregate demand (y_t, p_t), on the (constrained) level of output (y_t^c):

$$p_t^c = \frac{1}{\varepsilon}(y_t - y_t^c) + p_t \tag{2.19}$$

Subsequently, a monetary tightening, which pushes output of constrained firms downwards via the balance-sheet channel, and upwards via the input-price channel, has a simultaneous opposite effect on their prices via these channels (note that y_t^c and p_t^c are inversely related in the downward sloping demand curve (2.19)).

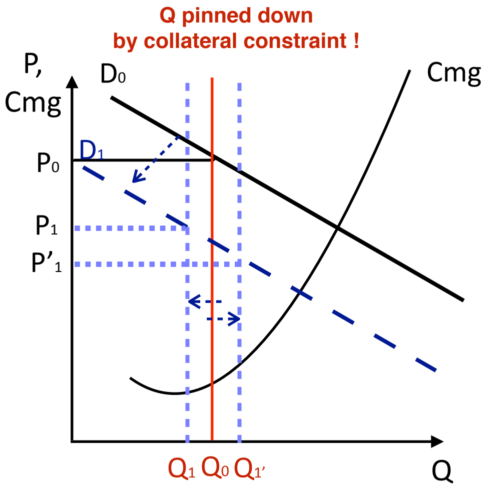


Figure 2.3: Transmission of a monetary tightening to a constrained firm

Finally, as shown by equation (2.20), monetary policy further affects future production levels of constrained firms via their investment decision ("investment channel") and via the associated debt level:

$$\frac{\partial E_t\{\widehat{y}_{t+1}^c\}}{\partial \varepsilon_t^m} = \overbrace{\frac{\partial \widehat{k}_{t+1}^c}{\partial \varepsilon_t^m}}^{\text{investment channel}} \overbrace{- \frac{(1-\alpha)\beta^{-1}(1-\theta^c)}{\nu - \beta^{-1}(1-\theta^c)} \left(\frac{\partial \widehat{\varrho}_t^k}{\partial \varepsilon_t^m} + \frac{\partial \widehat{i}_t}{\partial \varepsilon_t^m} \right)}^{\text{debt overhang}} \tag{2.20}$$

The sign and strength of these effects are directly linked to the ones on current output.

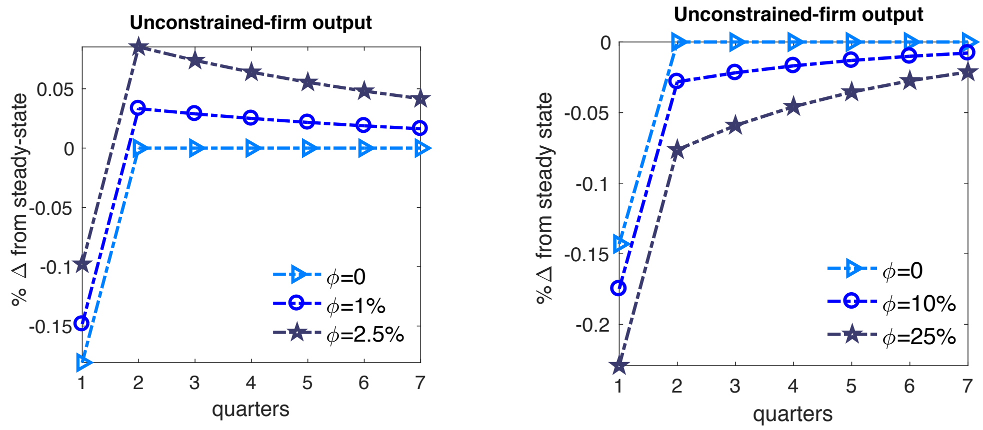
Figure 2.3 summarizes the transmission of a transitory monetary tightening to a constrained firm. On the one hand, it shows that the monetary impulse affects its output (Q_0) via the credit constraint and that the equilibrium effect can be of either sign, namely the vertical red line may shift to either of the two purple lines (i.e. to either Q_1 or Q_1'). On the other hand, it shows that monetary policy affects the price of a constrained firm (from P_0 to P_1 or P_1') via these effects and the shift engineered in its demand schedule (from D_0 to D_1).

Now, after having analyzed how monetary policy transmits to constrained firms, let's turn to how it transmits to *unconstrained ones*. On the one hand, monetary policy affects their output y_t^u and pricing decisions p_t^u by shifting their demand schedule:

$$y_t^u = -\varepsilon p_t^u + (\varepsilon p_t + y_t) \tag{2.21}$$

The demand for the goods of a monopolistic firm depends on the prices set by its competitions. Thus, in this heterogenous firm environment, the demand faced by unconstrained firms depends on the prices set by constrained firms.

Subsequently, monetary policy shifts the demand schedule of an unconstrained firm both *directly* via the consumption/saving decision of the household, and *indirectly* via its effects on the pricing (and production) decisions of constrained firms. Thus, when monetary policy pushes upwards or downwards the output of constrained firms, it simultaneously pushes the output and prices of unconstrained firms in the opposite direction. Furthermore, the choices made by unconstrained firms depend on the (equilibrium) effect of monetary policy on their marginal costs. These equilibrium effects depend also in the current heterogenous firm setup on the effects of monetary policy on constrained firms.



(a) Dominant balance-sheet channel

(b) Dominant input-price channel

Figure 2.4: Transmission spillovers and the response of an unconstrained firm to a monetary tightening

To get an intuition on how transmission spillovers from constrained firms affect the decisions of unconstrained firms, figure 2.4 shows the response of unconstrained firms to a rise in the policy rate as the share of constrained firms increases. It considers separately the case where the balance-sheet channel dominates (left panel), and the case where the input-price channel dominates (right panel). As previously shown, in response to a monetary tightening, a strong balance-sheet channel translates in (i) strong positive pressures on the prices of constrained firms, and (ii) strong negative pressures on input prices, and hence on marginal costs. Both these two indirect effects counteract the negative effects on unconstrained firms via the demand channel. As the share of constrained firms increases, these (counteracting) spillovers become stronger. Consistently, the left panel in figure 2.4 shows how the latter dampen the net equilibrium effect of monetary policy on unconstrained firms, and how, once strong enough, they switch the sign of the net effect in their direction. As shown in the right panel of figure 2.4, the opposite is true in the case of a dominant input-price channel, where the spillovers enhance instead the negative effect on unconstrained firms via the demand channel.

2.5.2 Firm-level transmission and structural parameters

We have seen that monetary policy has a differential effect on constrained and unconstrained firms. But how does this difference depend on the features of the economy? I now use a calibrated version of the model to show that the differential response of constrained and unconstrained firms increases in the tightness of financial constraints and in the level of nominal input price stickiness.

In the baseline calibration (table 2.1) non-financial parameters equal the textbook values in Gali (2015), the capital pledgeability ratio ν takes a value similar to the one in Iacoviello (2005), the fraction of constrained firms ϕ is set to its estimate in the UK data, and θ^c is such that around 20% of capital in the constrained set is financed by net worth. A time period in the model is one quarter. In all experiments, for ease of comparison with estimated dynamic responses, I consider the effect of a transitory monetary tightening of 25 basis points (figure 2.5).

I start studying how the tightness of credit constraints affects the transmission of monetary policy by looking at the effect of the capital pledgeability ratio ν. Figure 2.6 shows that the lower this ratio, the stronger the reduction in output, investment and working capital by constrained firms in response to a rise in the policy rate (navy lines in left panels *versus* purple lines in right panels). Otherwise stated, a lower pledgeability of capital implies a relatively weaker input price channel[25].

[25] As opposed to the current the model, the model in Iacoviello (2005) predicts a positive relation between capital tangibility and the effect of monetary policy. The difference comes from the consideration of working capital credit in the current analysis, alongside investment credit (Iacoviello (2005) considers only investment credit).

Table 2.1: Baseline calibration

	Parameter	Value
Non-financial parameters		
Intertemporal elasticity of substitution	σ	1
Discount factor	β	0.99
Inverse Frisch elasticity of labor supply	φ	5
Share of capital in total output	α	0.25
Output variety elasticity of substitution	ε	9
Price adjustment cost parameter	ξ	average κ as with Calvo for $\theta = 0.75$
Inflation coefficient Taylor rule	ϕ_π	1.5
Output coefficient Taylor rule	ϕ_y	0.5/4
Credit frictions		
Share of constrained firms	ϕ	0.2
Capital pledgeability ratio as collateral	ν	0.8
Fraction of capital financed by net worth	θ^c	0.23

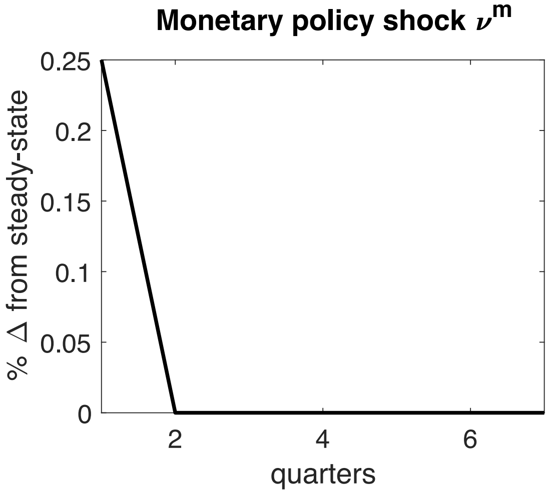
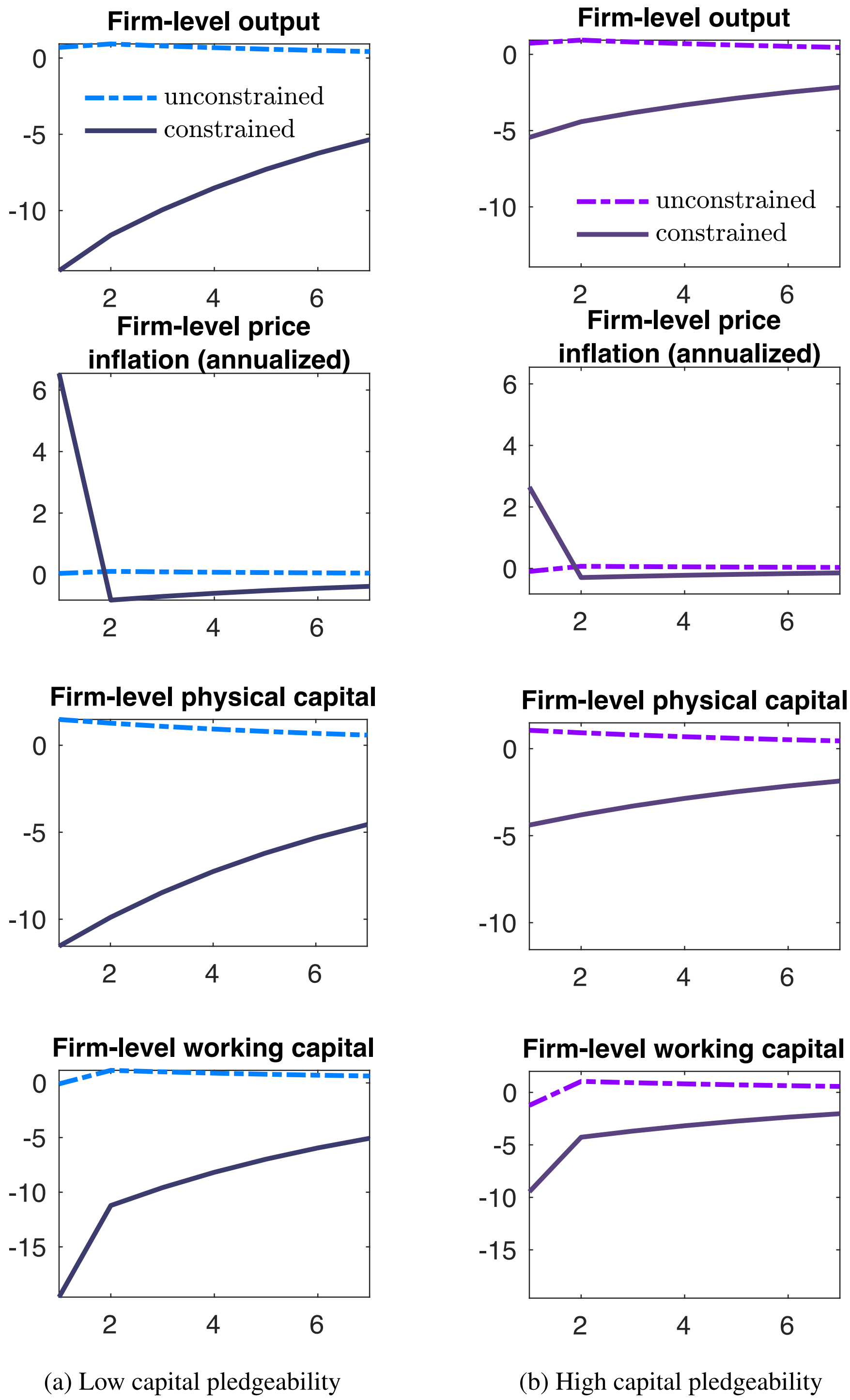


Figure 2.5: A transitory monetary tightening (25bp)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters



(a) Low capital pledgeability

(b) High capital pledgeability

Figure 2.6: Dynamic responses to a transitory 25 basis points monetary tightening
Note: Y-axis: % deviation from steady-state. X-axis: quarters

These results are explained by a decrease in the share of working capital credit to total credit as capital pledgeability ν decreases, as shown by its steady-state expression $1 - \frac{\beta^{-1}(1-\theta^c)}{\nu}$[26]. Thus, as capital becomes less plegeable, at given prices of labor and capital, the input-price channel gets relatively weaker, while the balance-sheet one gets relatively stronger. This is because constrained firms can finance to a lesser extent working capital credit against physical capital. Results are robust to changes in the share of constrained firms in the economy, and, as shown in figure 2.27 in the Appendix, also to an alternative calibration with a lower contribution of physical capital as production input as in Iacoviello (2005).

Furthermore, consistent with a stronger balance-sheet channel, given the downward-sloping demand schedule in equation (2.19), prices of constrained firms are steered more strongly upwards when capital is less pledgeable. This is because the stronger negative effect on constrained' firms output via the balance-sheet channel translates in a stronger positive effect on their prices. Also consistently, when capital pledgeability is low, transmission spillovers counteract to a larger extent the direct (negative) effects of a monetary tightening on unconstrained firms via the demand channel. Differences under baseline calibration are however small, and can be most easily noticed in the case of physical capital and working-capital.

Similar results are obtained when we consider the *net worth of constrained firms* (figure 2.28 in the Appendix). Specifically, firms with a lower net worth (and, hence, tighter credit constraints) reduce more strongly production and increase more their prices in response to a monetary tightening. Moreover, transmission spillovers counteract to a larger extent the negative effect of the monetary tightening on unconstrained firms via the demand channel. As shown by the steady-state share of working capital credit to total credit $1 - \frac{\beta^{-1}(1-\theta^c)}{\nu}$, these dynamics are also explained by a decrease in the ratio of working capital credit in total credit as net worth θ^c decreases.

Again same conclusions obtain under an alternative calibration with a lower contribution of physical capital as production input as in Iacoviello (2005) (figure 2.29 in the Appenidx). Interestingly also, under this alternative calibration, for high enough net worth, the relative strength of the input-price channel increases to the extent that it surpasses the one of the balance-sheet channel. Subsequently, output of constrained firms may be steered in a positive unconventional direction by a monetary tightening (solid navy line in figure 2.7). So, under the alternative calibration with $\alpha = 0.03$, there is a non-monotonic relation between firm new worth and the firm-level output response to monetary policy. Specifically, when the net worth of a firm is low, the firm is constrained and responds strongly to monetary policy. For higher levels of net worth, but low enough for the firm to remain constrained, the firm starts responding less and less to monetary policy,

[26]This ratio is computed using the expression of the collateral constraint (2.47) on page 84.

up to a point when it starts responding in an unconventional direction (region 1 in figure 2.8). This is because, as firm's net worth increases, the relatively strength of the input-price channel is enhanced, and ultimately it becomes stronger than the balance-sheet one. However, as its net worth increases and surpasses the level over which the firm becomes unconstrained, its output responds again in the conventional direction (region 2 in figure 2.8).

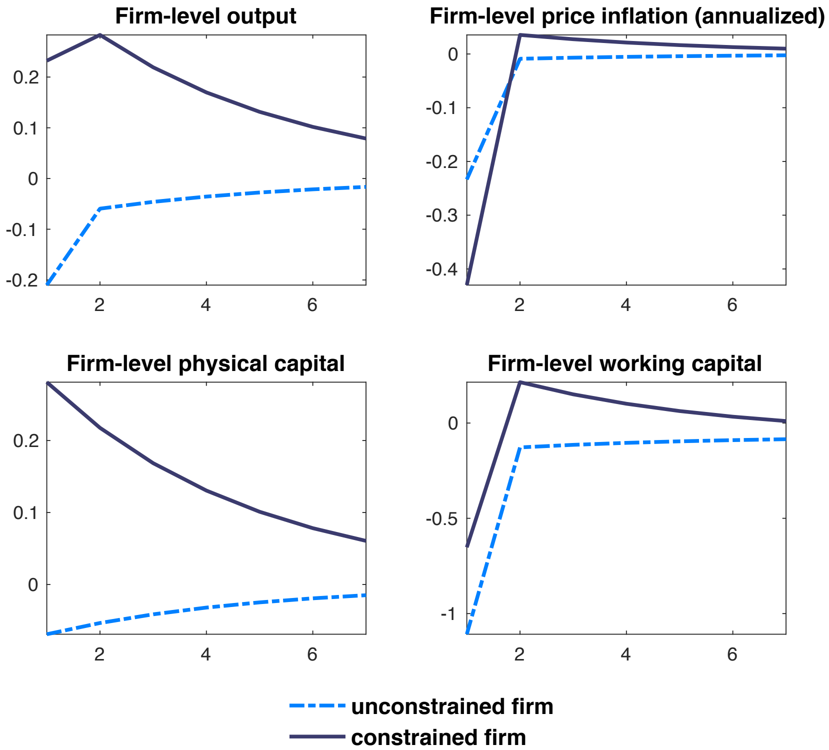


Figure 2.7: Dynamic response to a transitory monetary tightening (case with a relatively stronger input price channel: alternative calibration with $\alpha = 0.03$ as in Iacoviello (2005) and lower leverage $\theta^c = 0.45$)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

Finally, note in the right panel of the first row in figure 2.7 how in the case of a dominant input-price channel, constrained firms decrease their prices to accommodate the increase in output. Thus, in this particular case, spillovers push upwards the marginal costs of unconstrained firms and downwards their demand schedules, and hence reinforce the negative effects of the monetary tightening on unconstrained firms via the demand channel.

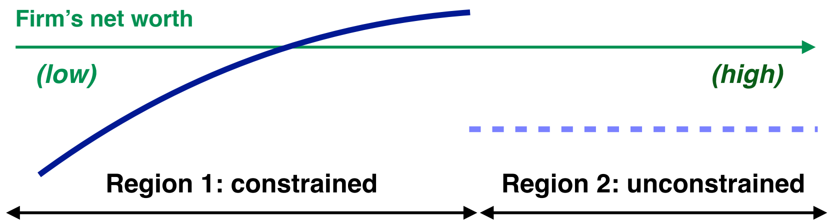


Figure 2.8: A firm's output response as a function of its net worth under the alternative calibration with $\alpha = 0.03$: navy solid line shows its response for values of net worth low enough for the firm to be constrained; light-blue dotted line shows its response for values of net worth high enough for the firm to be unconstrained

So far wages were assumed to be flexible. I now look at the effect of *wage stickiness* as a proxy (more generally) for input price stickiness. The sluggish adjustment of input prices is expected to weaken the relative strength of the input-price channel in equilibrium. To introduce wage stickiness, I follow the approach in Erceg, Henderson and Levin (2000). Specifically, I assume that the model economy is populated by a large number of identical households, each made up of a continuum of members specialized in a different labor service $j \in [0, 1]$. Household labor is now defined by an index of labor types:

$$L_t \equiv \int_0^1 \frac{L_t(j)^{1+\varphi}}{1+\varphi} dj \tag{2.22}$$

Labor decisions for each type j are taken at a union level with monopoly power over that labor type. Income is pooled within each household. The optimization problem of a typical household is identical to the one with flexible wages described in section 2.4.1, with the exception that L_t is now taken as given. Firms use all labor types and L_t represents the optimal mixture of these types. Otherwise, firms behave identically as in the flexible wage case[27].

In the presence of nominal wage rigidities, aggregate wage dynamics are described up to a first order approximation (in logs) by

$$\pi_t^w \approx \beta E_t\{\pi_{t+1}^w\} - \lambda_w \widehat{\mu}_t^w \tag{2.23}$$

where $\pi_t^w \equiv log\left(\frac{W_t}{W_{t-1}}\right)$ and $\widehat{\mu}_t^w \equiv \mu_t^w - \mu^w = \left(\omega_t - mrs_t\right) - \mu^w$ with $mrs_t \equiv \sigma c_t + \varphi l_t$ the economy's average marginal rate of substitution. Imperfect adjustment of nominal wages precludes real wages from moving one-for-one with households' average marginal rate of substitution. The wage inflation equation (2.23) replaces households' labor supply equation $\widehat{\omega}_t = \sigma\widehat{c}_t + \varphi\widehat{l_t}$ in the flexible

[27]For details see Gali (2015), Chapter 6.

wage case. Besides this modification, the model retains its baseline structure presented in section 2.4. To calibrate the new structural parameters, I follow Gali (2015), Chapter 6 and I set λ_w to match an average duration of wage spells of four quarters and, respectively, an average unemployment rate of 5% in the credit frictionless limit. A table reviewing the complete calibration is included in the Appendix on page 82. The credit-frictionless benchmark $\phi = 0$ is isomorphic (up to a first order approximation) to the one described in Gali (2015), Chapter 6.

When both types of firms populate the model economy, the nature of transmission mechanisms at the firm level is similar to the one with flexible wages. The sluggishness of nominal wage adjustments however directly weakens the input-price channel, and reinforces the strength of the balance-sheet channel in equilibrium. As a result, under baseline calibration, activity of constrained firms declines more when wages are sticky (right panels figure 2.30 in the Appendix) compared to when they are flexible (left panels in figure 2.30). Consistently, both the positive pressures on their prices and the spillover effects counteracting the negative impact of the monetary tightening on unconstrained firms are stronger[28].

2.5.3 Aggregate effect of monetary policy

I looked so far at how monetary policy transmits at the firm-level in an environment where some of the firms are financially-constrained. I now analyze how credit frictions shape the aggregate response of the economy to monetary policy. I have three new main findings with respect to existing literature. (i) Credit frictions do not necessarily amplify the response of aggregate activity to monetary policy. (ii) Observing constrained firms responding more than unconstrained ones to monetary policy does not necessarily imply that the aggregate response is also amplified by credit frictions. (iii) "Price puzzles" may emerge because monetary policy simultaneously shifts both aggregate demand and supply schedules of the economy.

The first two results arise because of the opposing nature of both transmission channels to the output of constrained firms, and of spillovers between the two types of firms. Aggregate amplification occurs in the case of a dominant balance-sheet channel when monetary policy affects more constrained firms than the unconstrained, only if (counteracting) spillovers to the unconstrained are weak. As shown in figure 2.9, this is the case under baseline calibration. Specifically, in this case, constrained firms reduce more production and investment than unconstrained ones in response to a rise in the interest rate (bottom panels), and, also, as the share of constrained firms (ϕ) increases, the effect of monetary policy on

[28]Same qualitative results obtain for the alternative calibration of $\alpha = 0.03$ proposed by Iacoviello (2005) (figure 2.31), as well as for different shares of constrained firms.

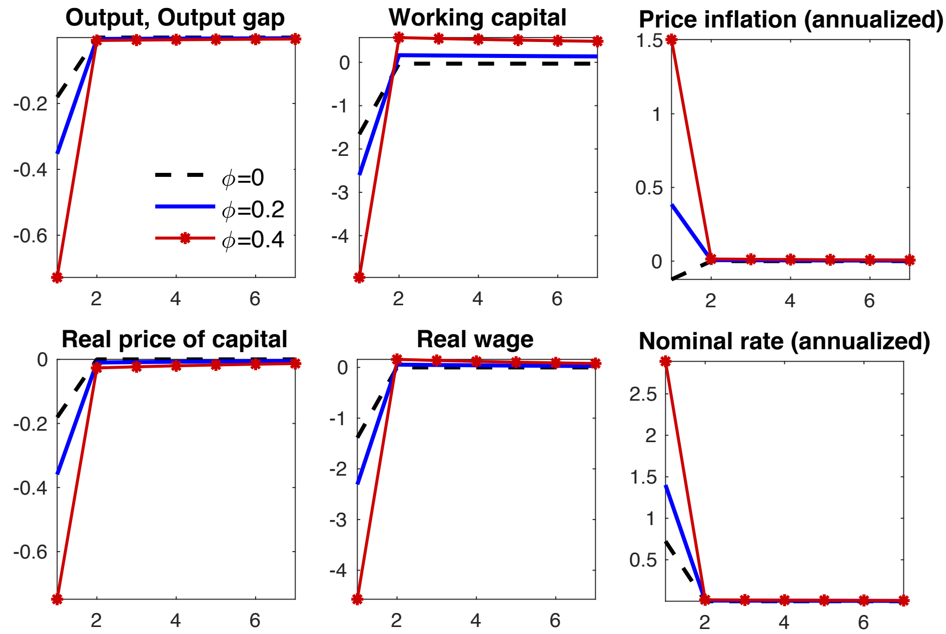
aggregate activity is amplified (top panels).

This is not however generally the case. Under the alternative calibration with $\alpha = 0.03$ as in Iacoviello (2005) (figure 2.10), because of general equilibrium (spillover) effects, even though constrained firms respond more than the unconstrained (bottom panels), aggregate responses are not (significantly) amplified as the share of constrained firms increases (top panels)[29]. One interesting implication of these results is that observing constrained firms responding more (on average) than the unconstrained (as observed in countries such as the UK or the US), does not necessarily imply that credit frictions amplify the response of aggregate activity to monetary policy (as usually concluded by empirical studies so far).

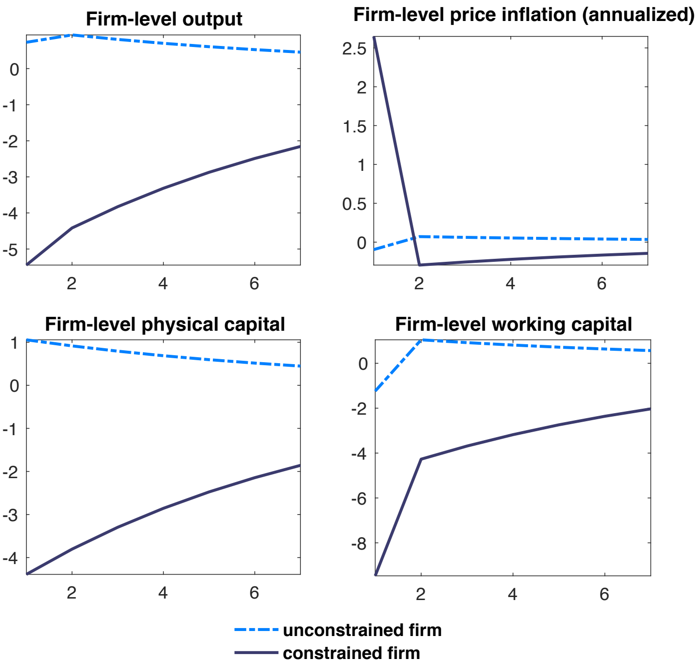
Finally, figure 2.11 further shows an example where credit frictions actually dampen the response of aggregate activity to monetary policy, namely where, as the share of constrained firms in the economy increases, aggregate activity declines less in response to a rise in the interest rate. This case characterizes the one with a dominant input price channel depicted in figure 2.7, where output of constrained firms is steered in an unconventional positive direction by a monetary tightening.

The third interesting finding in terms of aggregate monetary policy transmission is that a "price puzzle" may emerge when some of the firms in the economy face financial-frictions. Specifically, the aggregate price level may increase in response to a rise in the interest rate. In the basic version where all firms are financially-unconstrained price inflation always decreases. In the presence of credit frictions at the firm level however, if a monetary tightening reduces aggregate supply strongly enough (dotted red line in figure 2.12 (b)) together with aggregate demand (dotted black line in figure 2.12 (b)), the fall in aggregate activity may be associated in equilibrium with a rise in the price level. Such a case necessarily arises when the balance-sheet channel is dominant (figure 2.12 (a)). In this case, prices of constrained firms are pushed upwards by the monetary tightening. Thus, for high enough shares of constrained firms (e.g. larger than $\geq 20\%$ under baseline calibration), such positive pressures translate in an increase in the aggregate price level (figure 2.9, top panels). The bottom panels in figure 2.9 show how in this case the output of constrained firms is pushed strongly downwards, whereas their prices are pushed upwards.

[29]Moreover, note that for both such calibrations, the stronger response of investment by constrained firms is not associated to an amplification of aggregate investment since the latter is always equal to its credit frictionless counterpart.

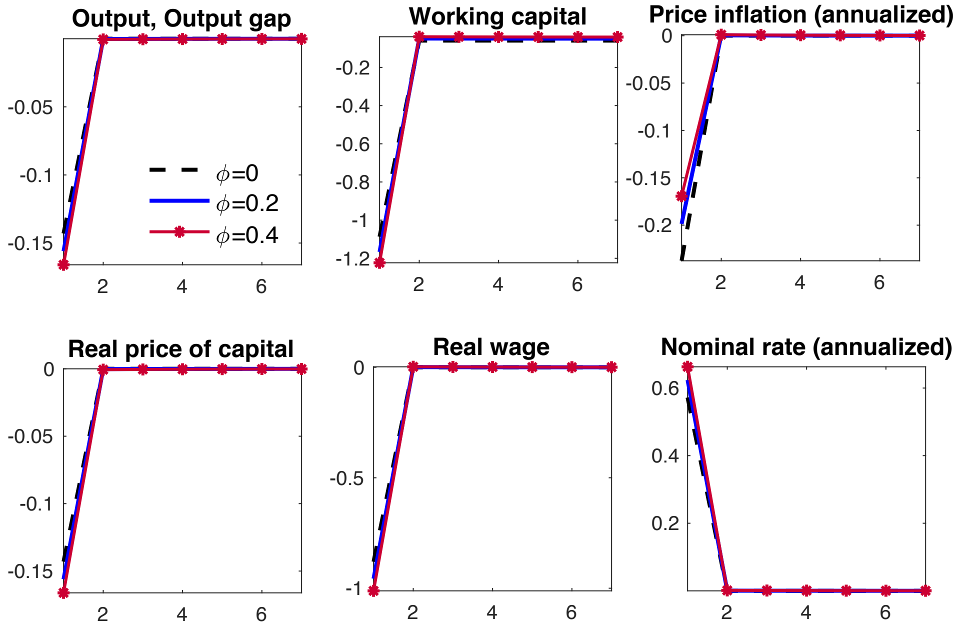


(a) Aggregate responses

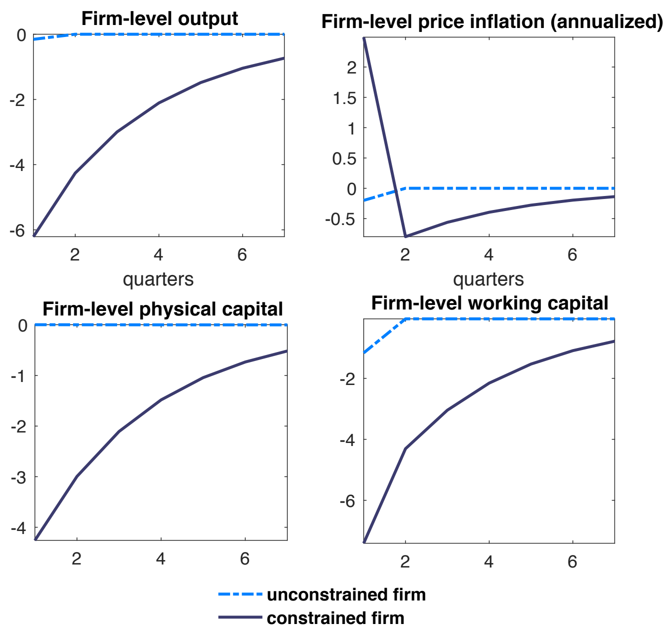


(b) Firm-level responses $\phi = 0.2$

Figure 2.9: Dynamic responses to a monetary tightening
<u>Note:</u> Y-axis: % deviation from steady-state. X-axis: quarters



(a) Aggregate responses



(b) Firm-level responses $\phi = 0.2$

Figure 2.10: Dynamic responses to a monetary tightening ($\alpha = 0.03$, Iacoviello (2005))
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

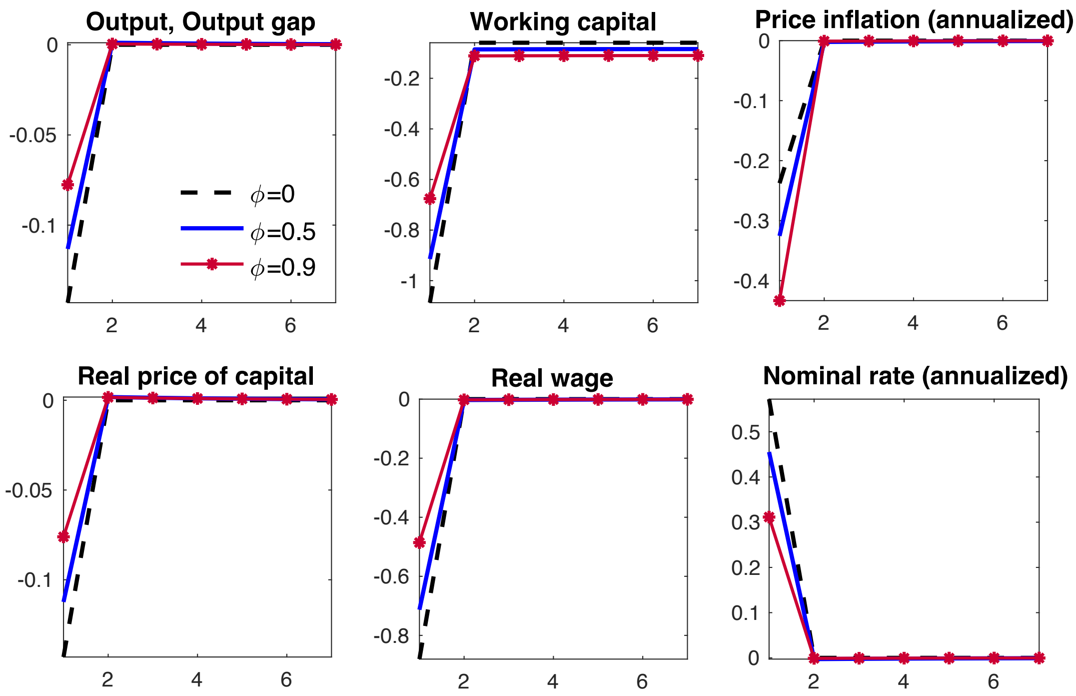
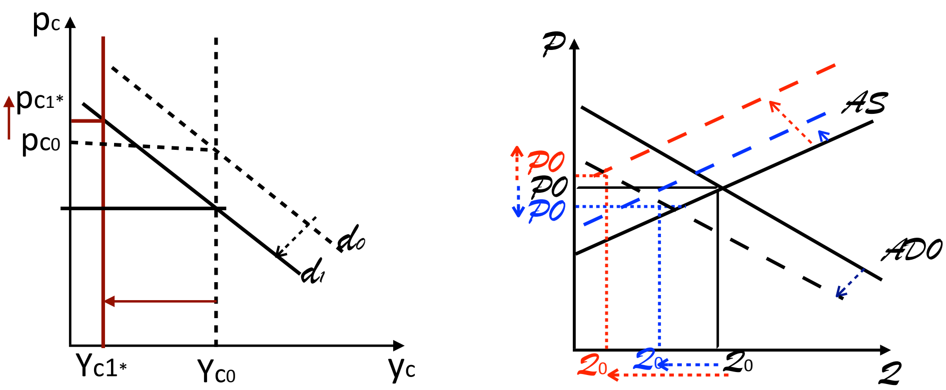


Figure 2.11: Dynamic aggregate response to a monetary tightening ($\alpha = 0.03$, $\theta^c = 45\%$)
<u>Note:</u> Y-axis: $\%$ deviation from steady-state. X-axis: quarters



(a) Firm-level

(b) Aggregate

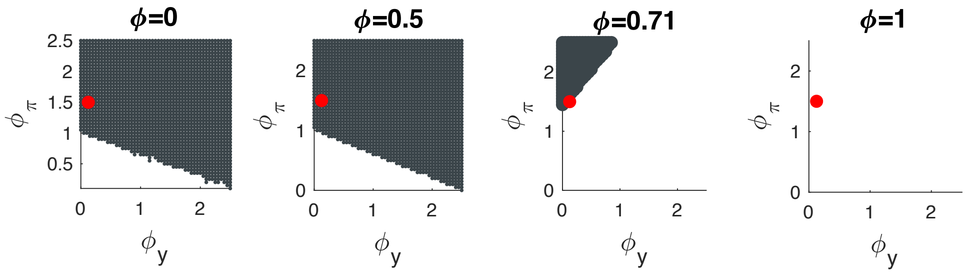
Figure 2.12: Strong balance sheet channel and the “price puzzle”

2.6 Monetary policy design

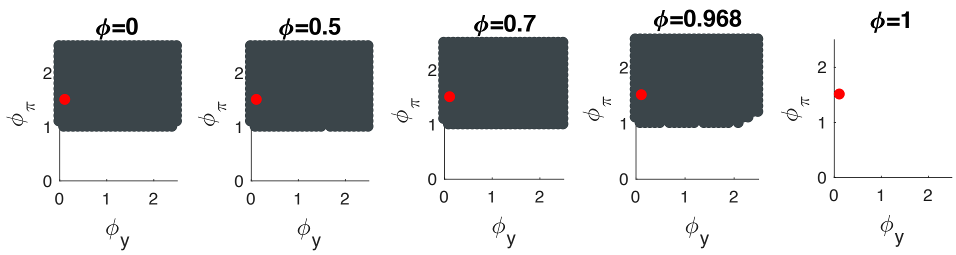
Since firms' credit frictions shape the effects of monetary policy, it is no surprise that they may ultimately affect as well the design of monetary policy. I first analyze how the financing frictions faced by firms may affect the stability properties of simple Taylor rules. I then focus on how monetary policy should take them optimally into account when deciding its response to business cycle fluctuations.

2.6.1 Equilibrium uniqueness and Taylor-rules

Given the changes in the transmission mechanisms of monetary policy, the properties that Taylor rules should satisfy to ensure equilibrium uniqueness may also be altered. I now study how these requirements change as the share of constrained firms increases. This question is important from a policy design perspective because, by ensuring that agents' expectations are anchored on an unique equilibrium path, the central bank avoids welfare losses due to sunspot fluctuations.



(a) Baseline calibration with sticky wages (dominant balance-sheet channel)



(b) Alternative calibration with $\alpha = 0.03$, $\theta^c = 0.45$ and flexible wages (dominant input-price channel)

Figure 2.13: Determinacy (in black) and indeterminacy (in white) regions
Note: Red dot indicates a standard Taylor rule with $\phi_\pi = 1.5$ and $\phi_y = 0.5/4$

According to the model, changes in the stability properties of simple Taylor rules reacting to price inflation and output are more likely to appear in an environment characterized by a strong balance-sheet channel. Specifically, under the baseline calibration with sticky wages characterized by a strong dominant balance-sheet channel such changes occur starting with a share of constrained firms in the economy around 70% (figure 2.13 (a)), whereas under the alternative calibration (i.e. with $\alpha = 0.03$ and $\theta^c = 0.45$) with flexible wages characterized by a dominant input-price channel they start occurring when this fraction is around 97% (figure 2.13 (b)).

2.6.2 Optimal monetary policy

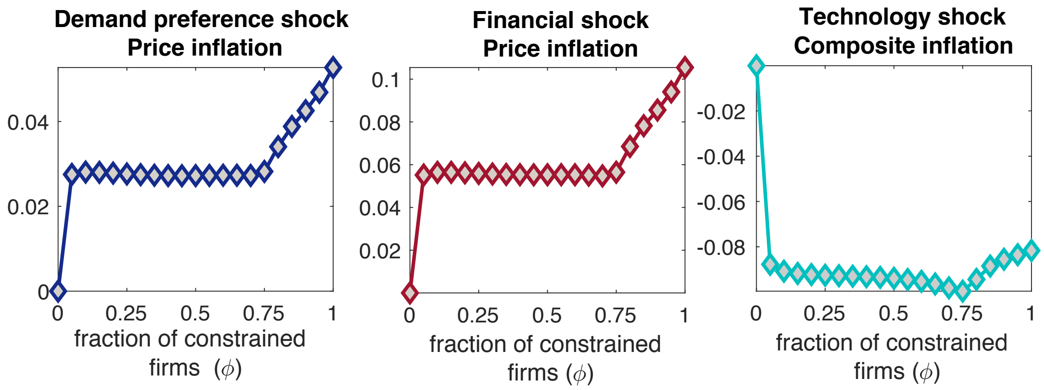
I now look how firms' financial constraints alter the optimal design of monetary policy. I assume employment is subsidized so as to correct for distortions associated to market power. Subsidies are financed by lump-sum taxes on households. The *flexible price* allocation in the credit-frictionless limit ($\phi = 0$) is thus the efficient benchmark at all dates (see page 82 in Appendix 2.11.2). I study optimal policy with commitment from a "timeless perspective" using the Linear-Quadratic approach. I focus on demand, technology and financial shocks. The latter is modeled as an exogenous variation in the collateral pledgeability ratio ν_t.

Welfare criterion On page 92 in the Appendix 2.11.5, I derive a second order approximation to households' discounted utility fluctuations around steady state. Associated welfare losses, expressed as a share of steady-state consumption, equal:

$$\begin{aligned}\mathcal{L}_0 \approx& \frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t\Big[\xi(1-\phi)(\pi_t^u)^2+\xi\phi(\pi_t^c)^2-(1-\sigma)\widehat{y}_t^2+\gamma^l\widehat{l}_t^2+\gamma^w(\widehat{\pi}_t^w)^2+\\ &\gamma^{lcu}\Big(\phi\frac{L^c}{L}(\widehat{l}_t^c)^2+(1-\phi)\frac{L^u}{L}(\widehat{l}_t^u)^2\Big)-\gamma^{yc}\widehat{y}_t^c-\gamma^{kc}\widehat{k}_t^c+\\ &+\gamma^{kcu}\Big(\frac{K^c}{K^u}\phi(\widehat{k}_t^c)^2+(1-\phi)(\widehat{k}_t^u)^2\Big)\Big]+t.i.p. \qquad (2.24)\end{aligned}$$

where $\gamma^l, \gamma^w, \gamma^{lcu}, \gamma^{yc}, \gamma^{kc}, \gamma^{kcu} \geq 0$ defined in the Appendix 2.11.5 (page 92), $t.i.p.$ terms independent of policy, and γ^{yc} and γ^{kc} are small due to small steady-state distortions (table 2.7 on page 95 in the Appendix 2.11.5).

I use the welfare criterion consistent with the model (2.24) together with the first-order approximation of the equations describing the functioning of the decentralized economy to derive the equilibrium dynamics under optimal monetary policy with commitment. Details are deferred to the Appendix 2.11.5 on page 92. I study the responses of the economy to demand, technology and financial shocks under optimal monetary policy, namely when the policy rate is chosen so as to maximize household's welfare. For brevity, given empirical results presented in the next section, I focus exclusively on cases with a dominant balance-sheet channel.



Note: Y-axis: Effect on impact of a positive transitory shock, % deviation from steady-state. X-axis: share of constrained firms (ϕ)

Figure 2.14: Departures from optimal policy regime in the credit-frictionless limit

Importance of spillovers Overall, transmission spillovers between constrained and unconstrained firms play an important role in the design of monetary policy. Specifically, optimal policy in the heterogenous case is not the simple weighted average of the one in the two polar cases (with only unconstrained and only constrained firms). Figure 2.14 (baseline calibration), shows that departures from the optimal monetary policy regime in the absence of credit frictions are relatively small unless the share of constrained firms is very large. A similar nonlinear pattern is obtained for the alternative calibration with $\alpha = 0.03$.

Optimal policy response to demand preference shocks Demand shocks affect only the efficient real interest rate, but not the efficient allocation. In the credit-frictionless limit (figure 2.15 for $\phi = 0$), monetary policy can replicate the efficient allocation by promising an aggressive response of the policy rate to variations in price inflation ("strict price inflation targeting"). Under this policy, whenever a shock z_t pushes aggregate demand (2.16) in one direction, monetary policy i_t offsets its effects by pushing it in the opposite one. Thus, in the absence of credit frictions, the central bank can perfectly insulate the economy from welfare losses caused by demand shocks.

Things are different when some of the firms in the economy are financially-constrained. First, their credit frictions distort the long-run allocation. Hence, the central bank has an incentive to engineer a positive variation in output irrespectively of the sign of the shock (i.e. engineer a positive variation in the output gap) so as to compensate for its long-run value being inefficiently low. Second, even if the central bank would like to insulate the allocation from the effects of such shocks, would not be able anymore. This is because, in contrast to the case without credit frictions, variations in the policy rate affect now directly both aggregate demand and aggregate supply. As a result, a variation in the policy rate aimed at offsetting the effects of the shock on aggregate demand has an additional supply-side effect. The latter induces a gap between the variation of the equilibrium allocation and the efficient one. As a result, in this case, demand shocks generally induce welfare losses under optimal policy, and the latter is conducted so as to minimize such losses. The nature of the optimal monetary policy regime, in particular its departure from price stability (the optimal regime in the absence of credit frictions), depends on structural parameters (e.g. figures 2.15 and 2.16).

When the balance-sheet channel dominates, a cut in the policy rate in response to a negative shock pushes up not only aggregate demand, but also aggregate supply via its positive effect on asset prices, and hence on collateral values. Thus, the cut in the policy rate that would offset the negative effects of the shock on aggregate demand, would also induce (inter-alia) a positive output gap and a decline in price inflation (because of additional supply-side effects).

Under the baseline calibration (figures 2.15) and the alternative one with $\alpha = 0.03$ (not shown), both characterized by strong balance-sheet channels, the positive supply-side effects of an interest rate cut are so strong, that the optimal response in the policy rate is very mild (and even slightly positive) so as to avoid strong deflationary pressures[30]. When the balance sheet channel is weaker (but still dominant), a policy rate cut has weaker negative supply-side effects on inflation. Hence, the policy rate declines more under optimal policy than in cases with a stronger balance sheet channel (figure 2.16)[31]. Importantly, because of the additional positive supply-side effects, the optimal decline in the policy rate is lower than in the absence of credit frictions. As a result, the optimal policy response is less likely to be constrained by the ZLB when there are constrained firms in the economy. Departures from strict price inflation targeting (the optimal policy regime in the absence of credit frictions) depend on structural parameters. In particular, a strong balance sheet channel (figure 2.15) implies small such departures, whereas a weaker one implies larger such departures (figure 2.16).

[30]Mirroring results are obtained for positive demand shocks (not shown).

[31]This is the sticky wage version of the second alternative calibration with $\alpha = 0.03$ and $\theta^c = 0.45$. With flexible wages, the real cost channel dominates under this calibration.

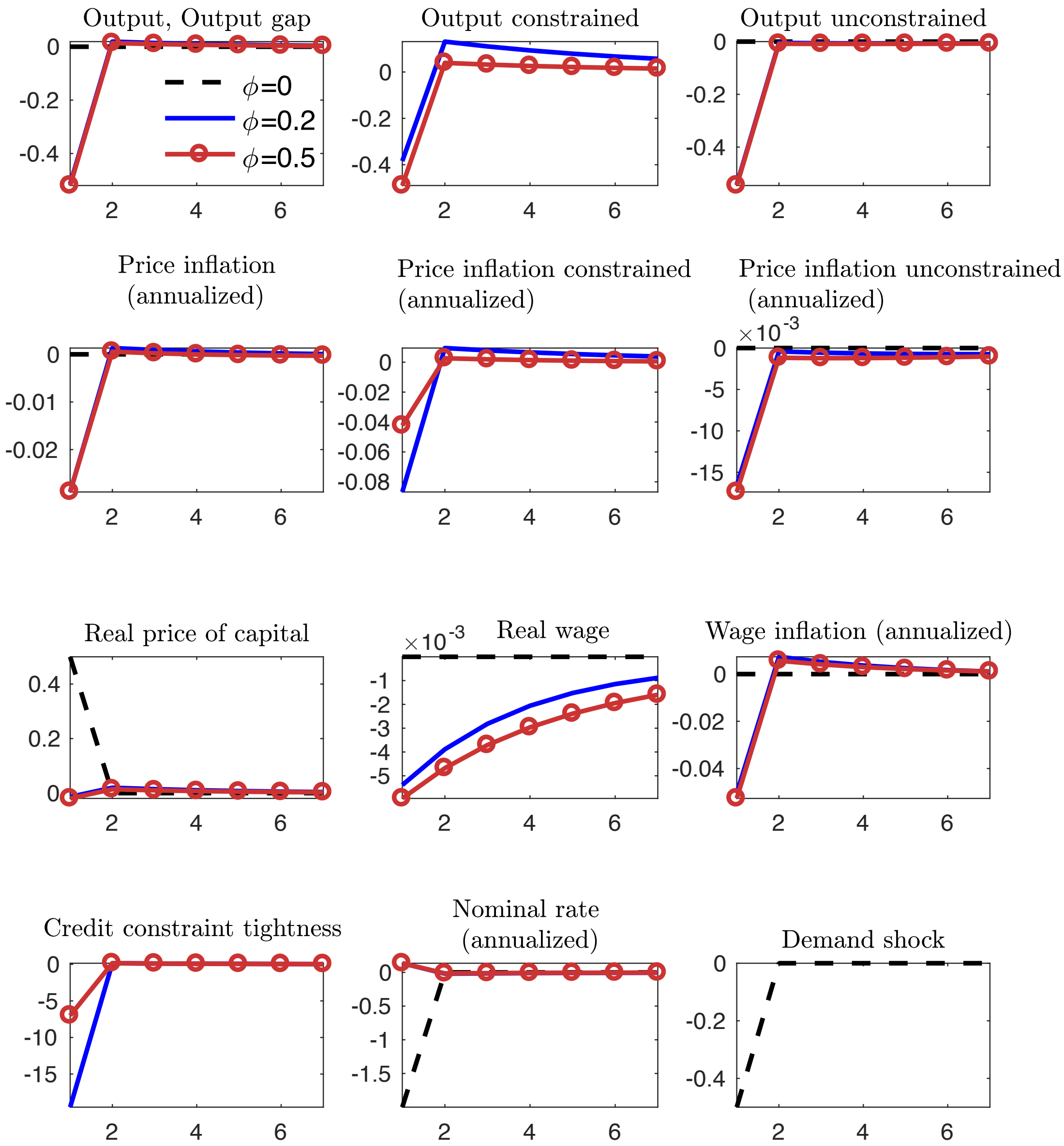


Figure 2.15: Optimal monetary policy: Dynamic responses to a transitory negative demand shock for different shares of constrained firms
<u>Note:</u> Y-axis: % deviation from steady-state. X-axis: quarters

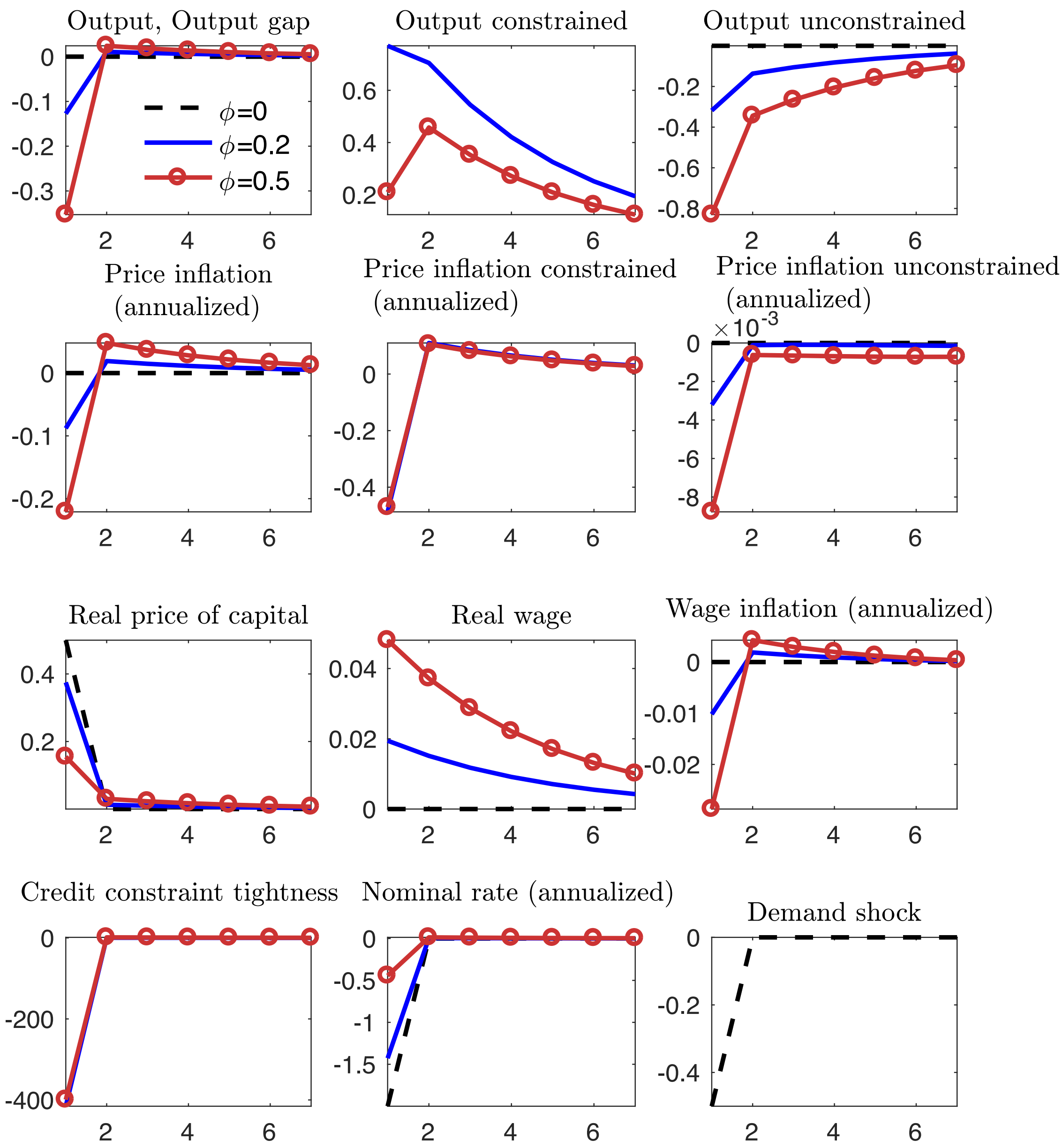


Figure 2.16: Optimal monetary policy: Dynamic responses to a transitory negative demand shock for different shares of constrained firms ($\alpha = 0.03$, $\theta^c = 0.45$)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

Optimal policy response to technology shocks Technology shocks induce a welfare-tradeoff for monetary policy even in the absence of credit frictions. Specifically, monetary policy can no longer simultaneously stabilize price inflation, wage inflation and close the gap between output and its efficient level (Galí (2015), Chapter 5). This is because stabilizing both wages and prices, and hence the real wage, is incompatible with the (efficient) variation of the latter needed to make output vary one-for-one with its efficient level.

In the absence of credit frictions, the optimal monetary policy response entails closing perfectly the output gap at the expense of variations in price and wage inflation (figure 2.17 for $\phi = 0$). The simple rule approximating well optimal policy requires responding to both price and wage inflation ("composite inflation targeting") with the strength of each response a function of the degree of nominal rigidities in goods and labor markets. In response to a positive technology shock, the interest rate declines so as to prop up demand, and hence so as to make up for the difficulties of firms to quickly reduce prices. In equilibrium, this increase in demand allows firms to produce at efficient levels (despite price stickiness).

Both the trade-offs and the nature of monetary policy transmission change in the presence firms facing credit frictions. In particular, in response to a positive technology shock, the monetary authority may now have to either cut the policy rate as in the credit-frictionless case, or to mildly increase it. As in the case of demand shocks, the result depends on the strength of deflationary pressures of positive supply-side effects associated to the decline in the policy rate. When such pressures are strong, as in the case of the baseline calibration or the alternative one with $\alpha = 0.03$, the central bank may have to mildly increase the policy rate– in this case, if the central bank were instead to decline it, collateral asset prices would go up allowing constrained firms to produce more, and hence amplifying deflationary pressures (figure 2.17). When the balance-sheet channel is weaker (i.e. under the calibration with $\alpha = 0.03$ and $\theta^c = 0.45$), deflationary supply-side effects are weaker, and the policy rate declines under optimal policy (figure 2.18). But, as for demand shocks, because of the additional supply-side effects, the policy rate cut is weaker under optimal policy than in the absence of credit frictions.

Optimal policy response to financial shocks Financial shocks become relevant for the design of monetary policy in the presence of firms' credit frictions. A negative shock to collateral pledgeability allows constrained firms to produce less, and given their downward sloping demand schedule, it pushes upwards their prices. Thus, such financial shocks act as cost-push shocks. This result has already been put forward in models with only constrained firms in the context of collateral constraints by Carlstrom et al. (2011), and of agency costs by De Fiore and Tristani (2013).

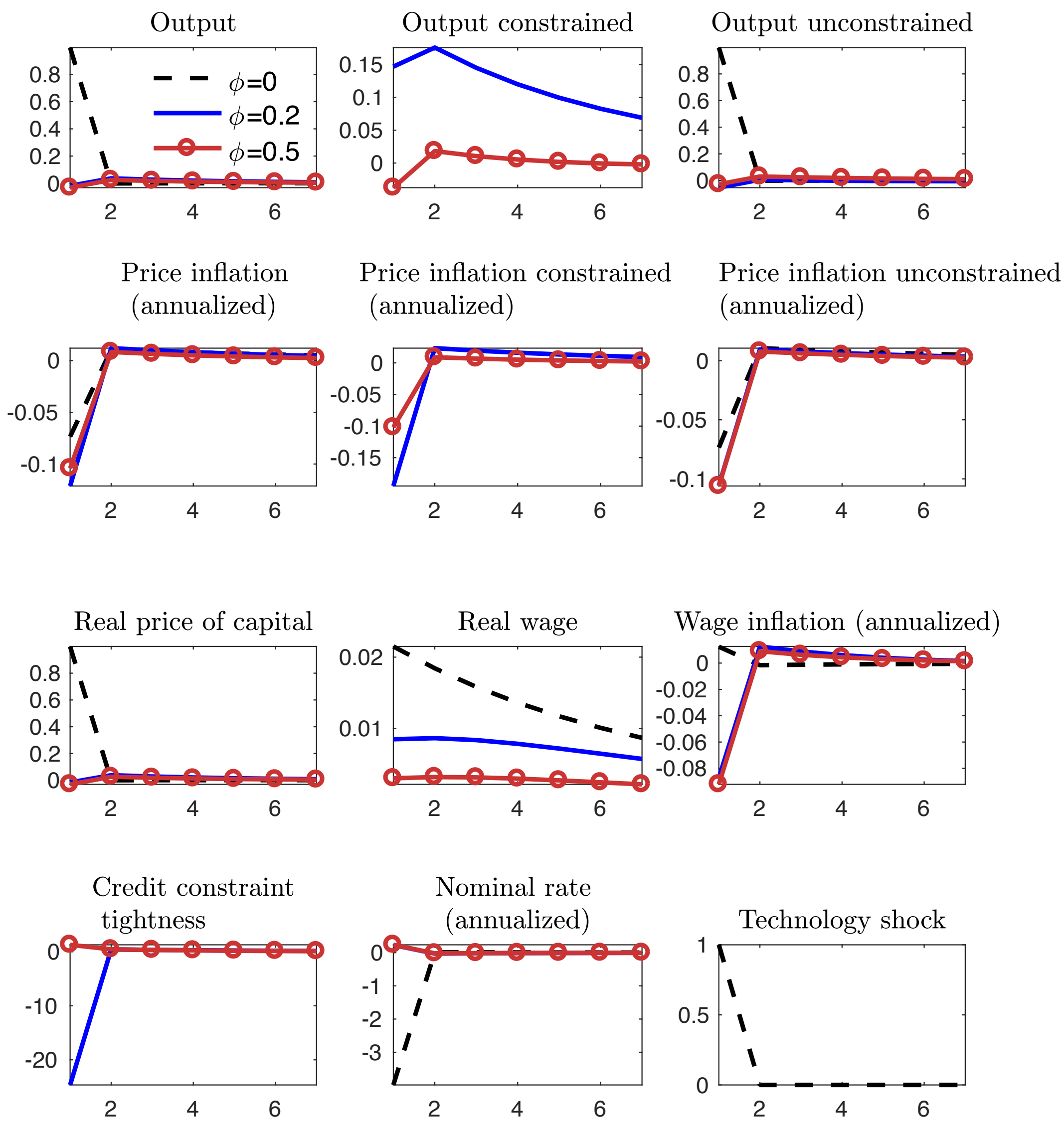


Figure 2.17: Optimal monetary policy: Dynamic responses to a transitory positive technology shock for different fractions of constrained firms
Note: Y-axis: % deviation from steady-state. X-axis: quarters

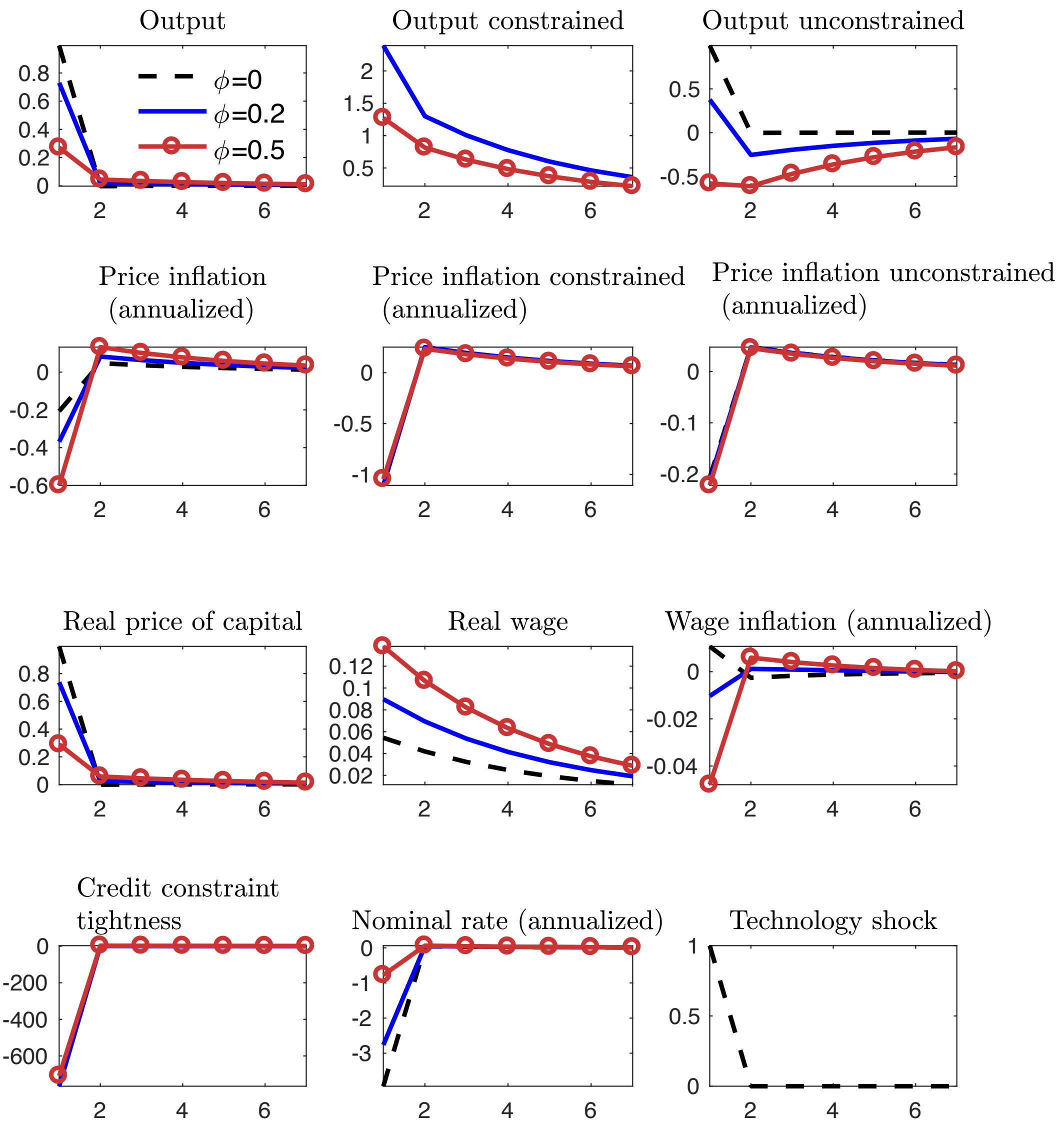


Figure 2.18: Optimal monetary policy: Dynamic responses to a transitory positive technology shock for different shares of constrained firms ($\alpha = 0.03$, $\theta^c = 0.45$)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

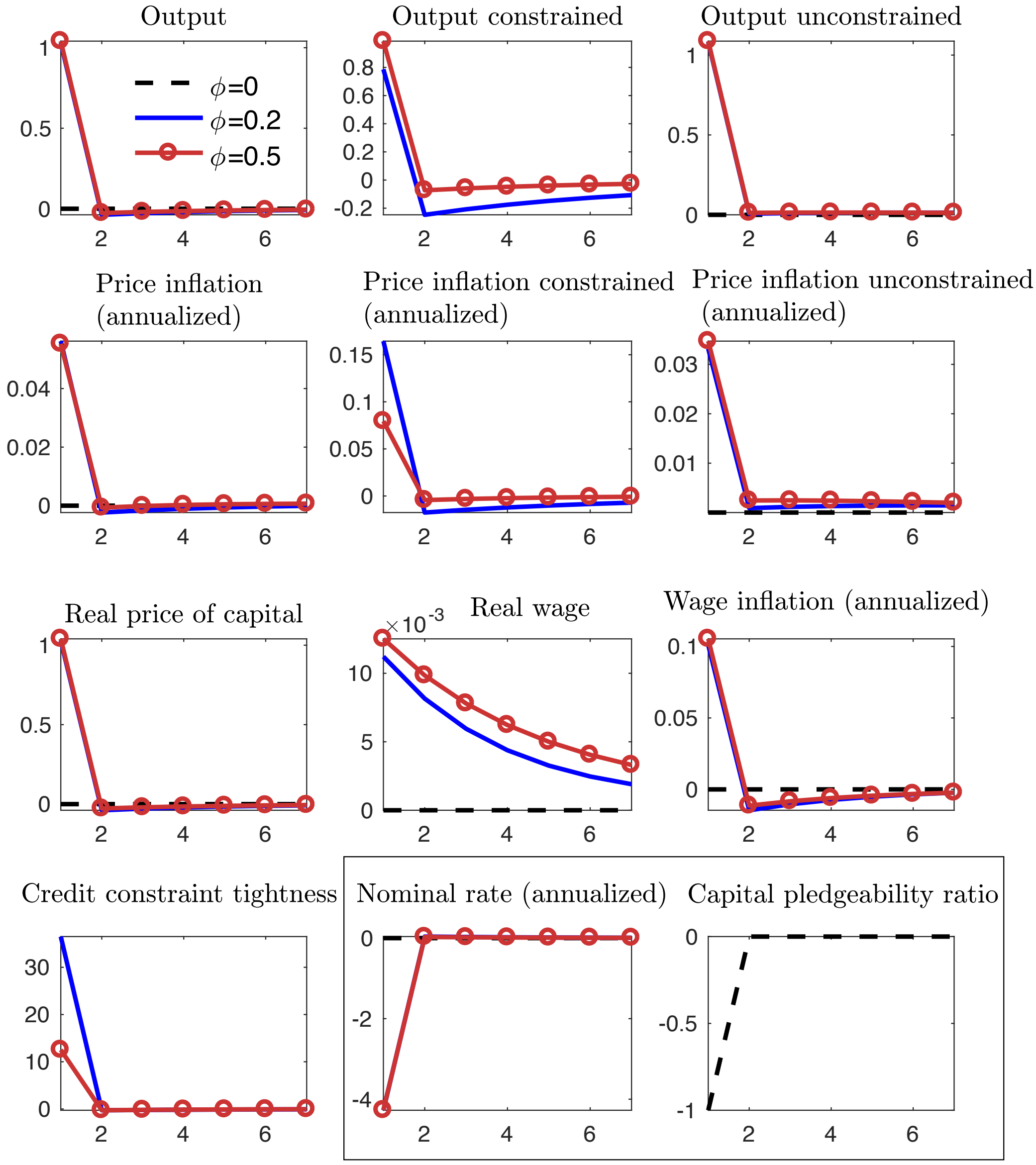


Figure 2.19: Optimal monetary policy: Dynamic responses to a transitory negative pledgeability ratio shock for different shares of constrained firms

<u>Note:</u> Y-axis: % deviation from steady-state. X-axis: quarters

As in Andres et al. (2013), but in contrast to Carlstrom et al. (2011), however, where collateral is not physical capital, the policy rate declines under optimal policy so as to prop up collateral asset prices, and hence to (partially) counteract the effect of the shock (figure 2.19). This is true under all three calibrations. Under the baseline calibration characterized by a strong dominant balance-sheet channel the decline is strong enough to actually increase collateral values in equilibrium. The latter effect allows constrained firms to produce above their steady-state levels and hence to contribute to the positive variation in the output gap. A similar result is obtained by De Fiore and Tristani (2013) in the context of costly-state verification. The latter however do not take into account the transmission spillovers between the two types of firms and the effect of steady-state distortions. Thus, relatively to these models, in the current heterogenous firm setup we can also see how the decline in the policy rate further induces unconstrained firms to expand production, and hence to push output above its inefficient steady-state level.

2.7 Empirical analysis

I now look whether the predictions of the model are in line with empirical evidence. In particular, I look whether the theoretical findings on firm-level transmission are corroborated by UK data. To do so, I extend the analysis in Cloyne et al. (2018)[32]. Specifically, relatively to this latter reference, I first study how monetary policy transmits to constrained firms depending on their characteristics (e.g. capital tangibility -as a proxy for pledgeability-, liquidity ratio). Cloyne et al. (2018) find that public UK firms with incorporation age less than 15 years and which do not distribute dividends reduce very strongly investment *on average* in response to a monetary tightening, whereas all others reduce it *on average* very weakly[33]. They also find that firm borrowing is highly correlated with collateral values, whereas the one of the latter group is not. They argue that these firms are most likely credit constrained, given that age is a good proxy for their track record in credit markets (hence, for credit access), and that not distributing dividends is a sign of a positive external finance premium (Fazzari et al. (1988), Mensa and Ljungqvist (2016), Jeenas (2018)). In the empirical analysis hereafter, I map the constrained set of firms in the model to this group, whereas the unconstrained one to all other firms in the sample.

[32]The empirical exercise is joint work with Ryan Banerjee (senior economist at the BIS).

[33]Their result is robust to controlling for size, asset growth, Tobin's Q, leverage or liquidity.

2.7.1 Methodology

The econometric methodology is based on an instrumental variable extension of the local projection method developed by Jorda et al. (2019). As in Cloyne et al. (2018) we use detailed financial statement data for publicly listed companies available from Thomson Reuters' WorldScope for the United Kingdom (table 2.2) and the five-year gilt yields for the interest rate[34]. The latter are instrumented with the series of monetary policy shocks constructed by Gerko and Rey (2017)[35]. These monetary shocks are obtained using the proxy-VAR/external instrument approach of Mertens and Ravn (2013) which uses movements in financial markets data (Short-Sterling Future contracts) in a short window around Bank of England policy rate announcements to isolate interest rate surprises. The sample spans from 1986 until 2018. Firms report for each fiscal year, but they do so in different months through the year. Thus, the data has a monthly dimension and refers to activity throughout the reporting year. All balance-sheet variables are converted to real values using the aggregate GVA deflator for the United Kingdom. For each observation, the interest rate is recorded at the end of the reporting month, and the series of monetary shocks used as instruments are also summed up to refer to the particular fiscal year. The asset tangibility ratio (used as a proxy for the capital pledgeability ratio in the analysis) is the ratio of tangible capital in total capital, where the latter is the sum of tangible capital and of an estimate of intangible capital computed using the methodology in Peters and Taylor (2017).

Baseline specification used to estimate the dynamic effects of a monetary tightening is a set of panel local projections of the form

$$X_{i,t+h} - X_{i,t-1} = \gamma_i^h + \sum_{g=1}^{G} \beta_g^h \cdot I[Z_{i,t-1} \in g] R_t + \sum_{g=1}^{G} \alpha_g^h \cdot I[Z_{i,t-1} \in g] + \varepsilon_{i,t+h} \tag{2.25}$$

with the dependent variable X the variable of interest (investment or working capital), h the number of (fiscal) years after the shock, $Z_{i,t-1}$ is a set of firm characteristics and the indicator function takes a value of 1 if firm characteristics fall in that particular firms's group. We include firm-fixed effects γ_i^h and monthly

[34]Variation in investment ratio winsored at 1%, variation in working capital at 5%, leverage at 1%.

[35]As in Cloyne et al. (2018), we drop firms within the finance, insurance, real estate and public administration sectors. Since the sample contains only public firms, to the extent that financial constraints are likely to be tighter for private firms than for public firms, the fraction of constrained firms in the UK economy, and hence the strength of spillover effects from constrained to unconstrained firms, are likely underestimated in our analysis. The use of monetary policy shocks ensures that the estimated effect of the variation in interest rate is not driven by other macro factors (namely, by endogenous components of monetary policy rule). We use the series of monetary policy shocks used in Cloyne et al. (2018) which was kindly sent to us by Paolo Surico.

Table 2.2: Firm-level balance-sheet data (WorldScope)

Variable ($X_{i,t}$)	Definition	WorldScope series
Investment	Investment rate = capital expenditures/ lag of property, plant and equipment	04601/ 02501
Working capital	log (working capital)	03151
Age	-	18273
Dividends paid	first lag	04551
Tangible capital	property, plant, equipment	02501
Leverage	total-debt/total-assets, first lag	03255/02999
Liquidity ratio	cash & short-term investments/total assets, first lag	02001/02999
Short term debt	first lag	03051
Total debt	first lag	03255

dummies[36]. The interest rate interacted with the dummies is instrumented by the monetary policy shock (also interacted with the dummies). The coefficients of interest are the β_g^h which measure how the effect of a monetary tightening on firm investment h (fiscal) years after the shock depends on firm's characteristics. As in the model, results are reported for a 25bps shock. Errors are clustered at the firm-level.

2.7.2 Results

We start by running the set of regressions for the two groups "constrained" and "unconstrained" for investment (as in the original paper) and working capital. The β_h^g coefficient for the two groups at different horizons are plotted in figure 2.20. Firms in the constrained group are found to reduce their working capital and investment more than unconstrained ones in response to a monetary tightening. Through the lenses of the model, this implies that in the UK monetary policy affects the activity of constrained firms relatively stronger via the balance-sheet and debt-deflation channels.

Now we try to understand transmission within the constrained set in more detail, by looking at the effect of parameters analyzed in the theoretical section: capital tangibility and liquidity ratio ("net worth"). We also try to identify instances where the activity of constrained firms is steered in a positive unconventional direction by a monetary tightening (i.e. a stronger input-price channel). We first explore how the responses of constrained firms to monetary policy depend on the

[36]Firm fixed effects capture permanent differences in investment behaviour across firms.

pledgeability of their physical capital as collateral. As more tangible capital is expected to be more pledgeable as collateral, we use the former as a proxy in the empirical analysis. We start by splitting the constrained set at each date in two groups based upon their capital tangibility ratio in the previous fiscal year. The set of regressions (2.25) is thus specified based on the following three groups g: constrained firms with high capital tangibility, constrained firm with low capital tangibility and unconstrained firms.

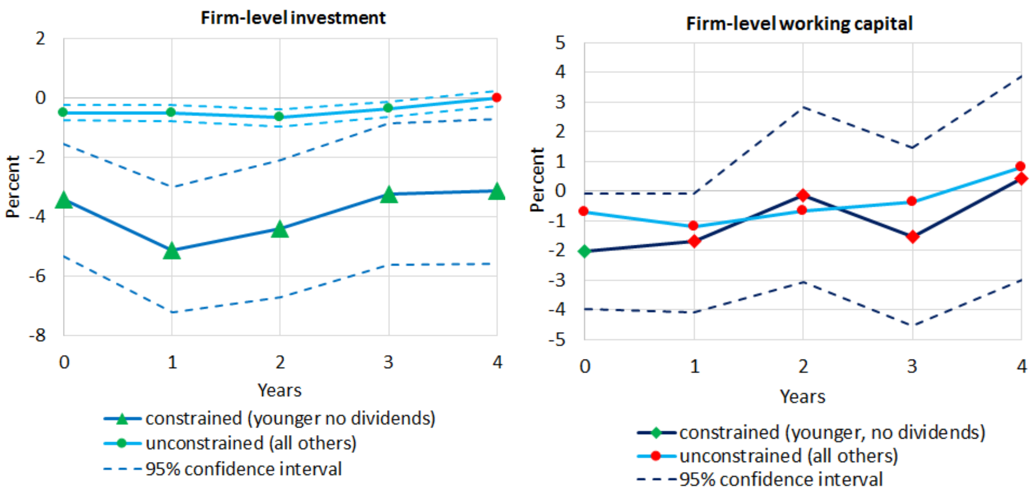
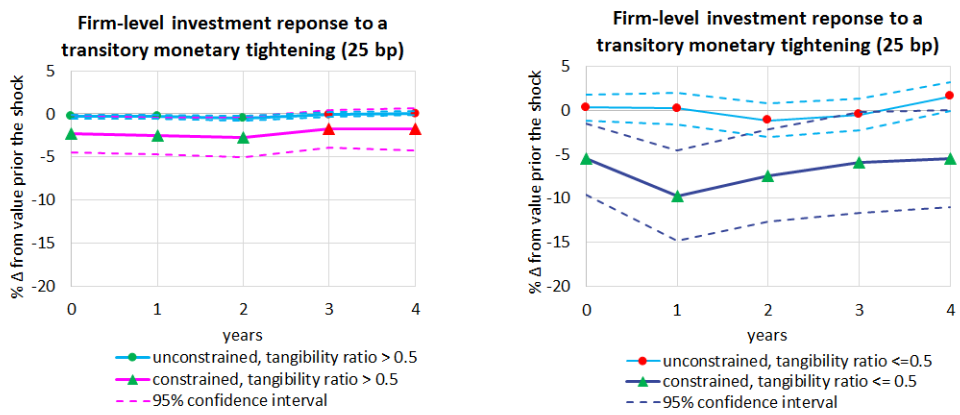


Figure 2.20: Estimated responses to a monetary tightening in the UK

Figure 2.21 plots the β_g^h coefficients in the case of investment for the two subgroups in the constrained set and for the unconstrained group. The left panel compares the response of constrained firms with high capital tangibility (purple line) with the one of unconstrained ones (light blue), whereas the right panel compares the response of constrained firms with low capital tangibility (navy line) with the one of unconstrained firms. The figure shows that firms in the constrained group with low capital tangibility respond the most. If not paying dividends is a good proxy for being financially constrained (and hence, assuming that all firms in this group are constrained is a good approximation), this result implies that the strength of the balance sheet channel decreases with the pledgeability of capital.

Table 2.3 shows that these results hold more generally. Specifically, it shows the results of a regression with "constrained" and "unconstrained" firms as the two groups g, and an interaction term between being in the constrained group and the capital tangibility ratio in the fiscal year prior to the shock. The sign of the latter coefficient of the interaction term is positive and significant, implying that a higher capital tangibility ratio is associated to a weaker decline in investment, and hence to a relatively weaker balance-sheet channel. This result is in line with the theoretical predictions discussed in the previous section (figure 2.27 on page 87 the Appendix). In the model this result arises because a decrease in the pledgebility of capital implies a decrease in the share of working capital credit in total credit, and hence of a decrease in the relative strength of the real cost channel. Importantly, a model with only investment credit as the one in Iacoviello (2005) would imply the opposite. Thus, our empirical results may also be taken as evidence of the importance of working capital credit for constrained firms in the sample.



(a) High tangibility

(b) Low tangibility

Figure 2.21: Estimated investment responses to a monetary tightening (25bp)

Table 2.3: Effect of the asset tangibility ratio on the magnitude of the response of credit constrained firms to monetary policy in the UK

	Investment ratio				
	(h=0)	(h=1)	(h=2)	(h=3)	(h=4)
r5 Indy	-5.925 ***	-11.003***	-9.535***	-7.068***	-6,00***
r5 Indytarlag	4.067**	9.13***	8.401***	6.038***	4.4326 **
r5 Iu	-0.506***	-0.489***	-0.659***	-0.375**	-0.025

Regression: $I_{i,t+h} - I_{i,t-1} = \gamma_i^h + \beta_{indy}^h \cdot Indy \cdot R_t + \beta_{indyta}^h \cdot Indytarlag \cdot R_t + \beta_{iu}^h \cdot Iu \cdot R_t + \alpha_{indy}^h \cdot Indy + \alpha_{indyta}^h \cdot Indytarlag + \alpha_{iu}^h \cdot Iu + \alpha_{indy}^h \cdot Indy + i.m + \varepsilon_{i,t+h}$
where $Indy$ is an indicator variable which equals 1 if the firm is younger than 15 years and does not distribute dividends, $Indytarlag$ is an interaction variable with the (lagged) tangibility ratio of firm's assets and Iu is an indicator variable which takes the value 1 when $Indy = 0$.
The interest rate is instrumented with the series of monetary policy shocks.

* $p < .1$, ** $p < .05$, *** $p < .01$

Also in line with theoretical predictions is the positive highly significant correlation between capital tangibility and leverage in the constrained group shown in table 2.4. This correlation implies that, as in the model, constrained firms with low capital pledgeability tend to be less levered[37]. Furthermore, consistent with theoretical predictions, there is no such significant correlation in the group of old firms which distribute dividends (the most likely "unconstrained" group).

Table 2.4: Correlation asset tangibility ratio and leverage (significance level)

constrained (younger-no dividends)	**0.1037** (0)
older-dividends (i)	0.001(0.8976)

* $p < .1$, ** $p < .05$, *** $p < .01$

We now look how the response of constrained firms to monetary policy depends on their ability to finance production with their own net worth, otherwise stated how the tightness of their credit constraint affects the response to monetary policy[38]. We use the liquidity available prior to the shock as a proxy for the funds that the firm could use to finance production aside credit. As in Cloyne et al. (2018) we normalize liquidity by the size of total assets. We split the constrained set at each date in two groups based upon their liquidity ratio in the previous fiscal year. The set of regressions (2.25) is thus specified based on three subgroups g: constrained firms with high liquidity ratio, constrained firm with low liquidity ratio and "unconstrained" firms. We use 20% to define the threshold between low and high liquidity ratios[39]. Figure 2.22 plots the β_g^h coefficients in the case of working capital for the two subgroups in the constrained set. As in the model, firms that can use less of their own funds to finance production (here, firms with a low liquidity ratio) respond more to monetary policy.

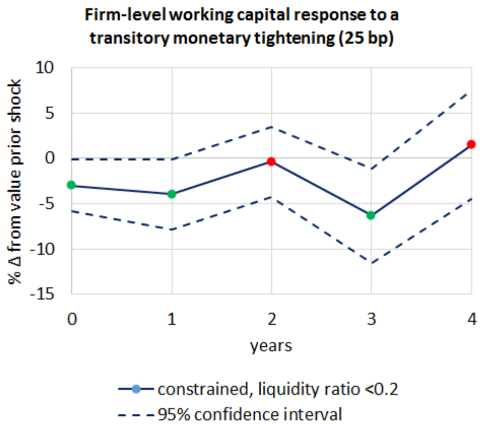
In the theoretical analysis we saw that a strong input price channel is favored by a high share of working capital credit in total credit. We do not have information on working capital credit, but we do on short-term debt and we use the latter as a proxy. We look at the limiting case of firms with only short-term debt and focus on the ones with a low pledgeability of their capital. Specifically, we define two groups g one including the latter group of firms, and the other all firms in the economy and run the regression in (2.25). The β^g coefficients are plotted in

[37]In the model, lower capital pledgeability is associated to lower steady-state leverage ratios of constrained firms.

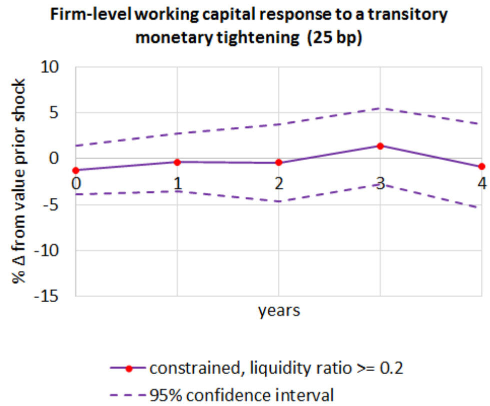
[38]Constrained firms with higher net worth need less credit, and hence their financial constraint is expected to be looser.

[39]Around $2/3$ of observations are below this threshold. Results are robust and differences between the two subgroups more striking when defining a lower threshold of 10% which splits the constrained group more evenly in two sub-groups.

the left panel of figure 2.23. The ones for the constrained groups with short-term debt only and capital pledgeability ratio lower than 0.9 are plotted in navy blue, whereas the ones for all other firms in light blue[40]. The figure shows that investment by firms in the first group is steered by a monetary tightening in a positive unconventional direction on impact in line with a relatively stronger input-price channel. The right panel of figure 2.23 shows that the model makes a similar prediction in this special case. Specifically, both the theoretical responses of investment and output are steered in an unconventional positive direction when the firm has only working capital credit and a low capital pledgeability ratio[41].



(a) Low liquidity-ratio



(b) High liquidity-ratio

Figure 2.22: Estimated working capital responses to a monetary tightening (25bp)

More generally, by running the regression in (2.25) with two groups of firms constrained and unconstrained and an interaction term between being constrained and the lagged value of short term to long term ratio, we obtain a positive significant coefficient for the latter (table 2.5). This implies that, to the extent that short term debt is a good proxy for working capital credit, as in the model, a higher share of short term debt is associated to a weaker response of a constrained firm to a monetary tightening[42].

[40]Initially, I run a regression with four groups:. However, results for the constrained group with only short term debt and low capital pledgeability are the same as for the simpler regression.

[41]The pledgeability ratio is set low enough for a set of firms to be constrained in equilibrium under the alternative calibration following Iacoviello (2005).

[42]Regression: $I_{i,t+h} - I_{i,t-1} = \gamma_i^h + \beta_{indy}^h \cdot Indy \cdot R_t + \beta_{indyta}^h \cdot Indytarlag \cdot R_t + \beta_{iu}^h \cdot Iu \cdot R_t + \alpha_{indy}^h \cdot Indy + \alpha_{indyta}^h \cdot Indytarlag + \alpha_{iu}^h \cdot Iu + \alpha_{indy}^h \cdot Indy + i.m + \varepsilon_{i,t+h}$ where $Indy$ is an indicator variable which equals 1 if the firm is younger than 15 years and does not distribute dividends, $Indytarlag$ is an interaction variable with the (lagged) tangibility ratio of firm's assets and Iu is an indicator variable which takes the value 1 when $Indy = 0$. The interest

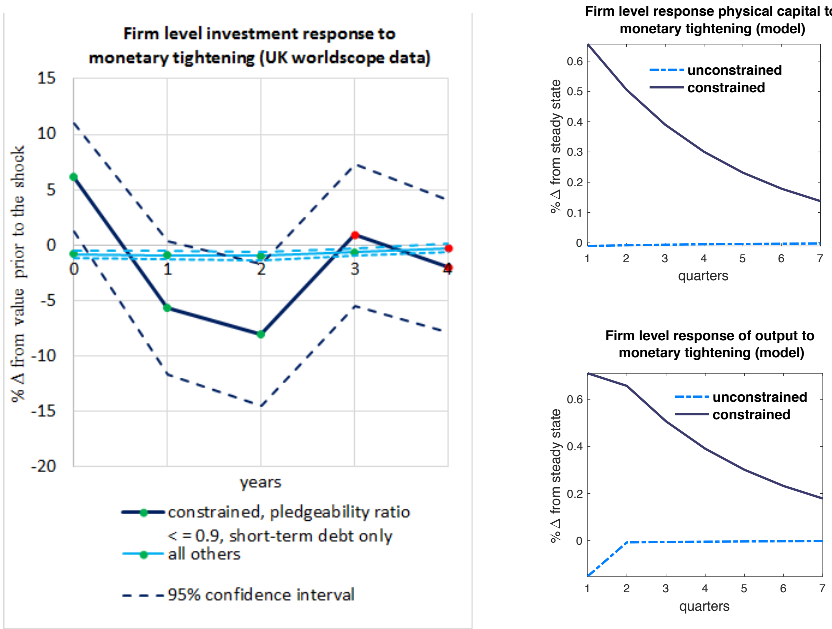


Figure 2.23: Estimated investment responses to a monetary tightening (left). Theoretical investment responses to a monetary tightening; alternative calibration 2 with flexible wages, $\theta^c = 1$ i.e. 100% of physical capital financed by equity and $\nu = 0.2$ i.e. 20% of physical capital is pledgeable as collateral (right)

Table 2.5: Share of short-term debt in total debt $sh/lt - lag$ and the response of credit constrained firms to monetary policy in the UK

	Investment ratio (h=0)
r5 Indy	-5.037***
r5 $Indy - sh/lt - lag$	4.81**
r5 Iu	-0.477***

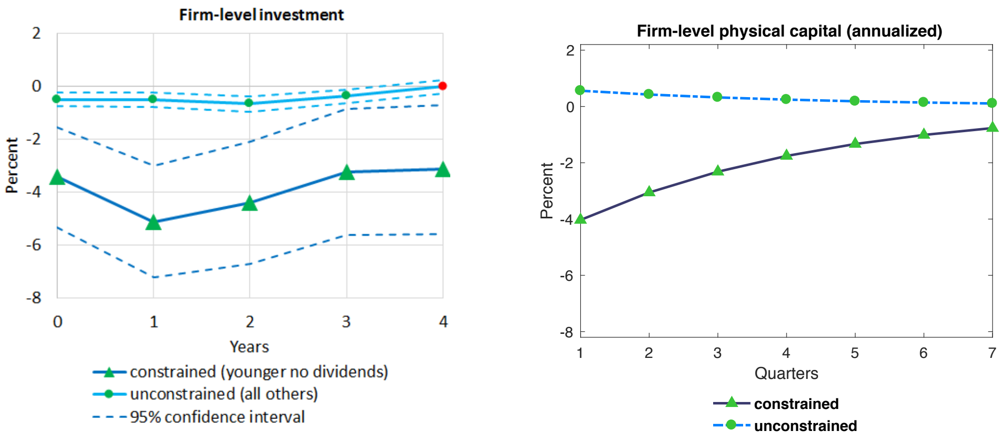
* $p < .1$, ** $p < .05$, *** $p < .01$

rate is instrumented with the series of monetary policy shocks as described in table 2.2.

2.8 Use of the model for policy analysis

The model can be used to give a structural interpretation to the estimated dynamic responses to a monetary tightening shown in the previous section. According to those estimations (reported again in the left panel of figure 2.24), financially-constrained firms reduce their activity stronger than unconstrained ones. By setting the capital leverage ratio of constrained firms at $\theta^c = 0.32$ under the alternative calibration with $\alpha = 0.03$, we can match reasonably well the empirical firm-level responses (right panel of figure 2.24).

The empirical results can be explained through the lenses of the model by a dominant balance-sheet channel. Specifically, as the central bank rises the policy rate, collateral asset values decline, and hence constrained firms are compelled to cut strongly production. The strong adverse effects on constrained firms further spill over to unconstrained ones, benefiting the latter. Specifically, the strong decline in production by constrained firms decreases competition on output markets for the unconstrained ones, and implies strong negative pressures on input prices (because of the strong decline in input demand). These positive spillover effects via output and input markets partially counteract the negative effects of the monetary tightening on the firms which do not face financing constraints. Hence, they respond only mildly in equilibrium.



(a) Estimated responses (UK) (b) Theoretical responses

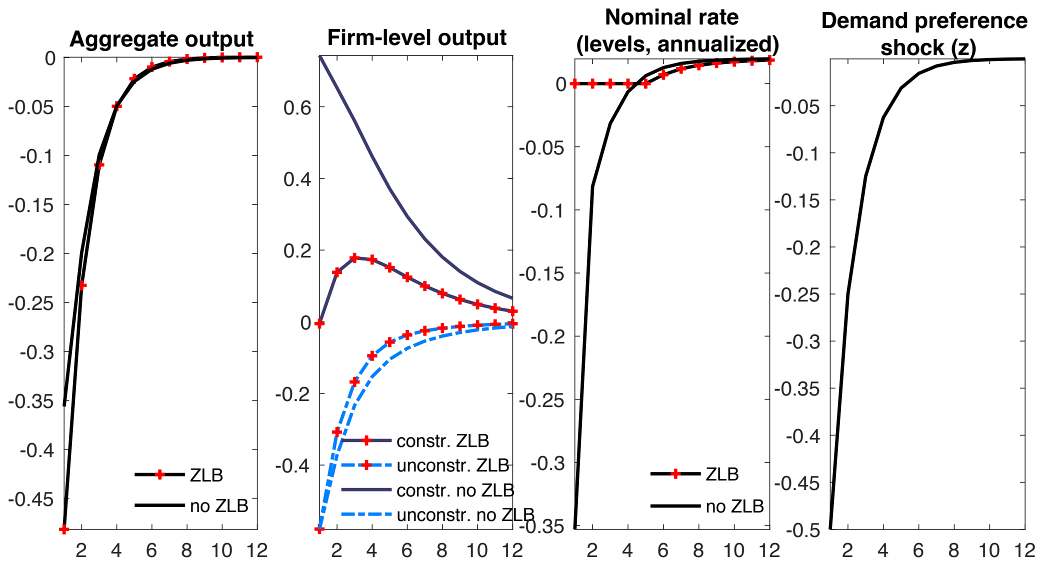
Figure 2.24: Response of investment to a monetary tightening (25 bps)

The model may also help us understand better how monetary policy transmits in the current low interest rate environment where the probability that the policy rate hits the ZLB is high. In particular, the model outlines that in economies with strong dominant balance-sheet channels (such as the UK, the US, and most likely the Euro Area), the presence of an effective ZLB on the policy rate particularly hurts financially-constrained firms, while unconstrained ones may benefit. To see why, note that in the absence of the ZLB, constrained firms are more negatively affected than unconstrained ones in response to a monetary tightening, but more positively affected in response to a monetary loosening. Thus, in the absence of the ZLB, the relatively gains and losses for these firms compensate over the business cycle.

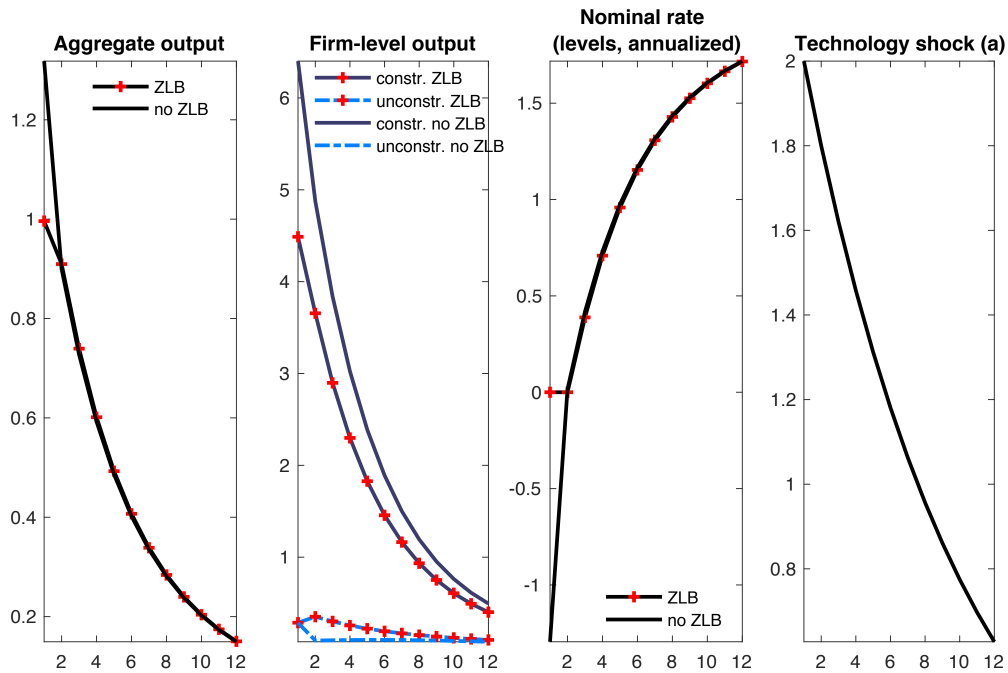
Since the ZLB limits the more positive effects of monetary policy when an interest rate cut is warranted, but the more adverse effects remain unchanged, on average over the business cycle, constrained firms end up particularly hurt. Furthermore, because of spillover effects, unconstrained firms benefit. Specifically, since constrained firms are affected less positively by a cut in interest rates, the associated negative spillovers to the unconstrained ones are also lower. As a result, the net positive effects of monetary policy on unconstrained firms in times where a monetary loosening is warranted are stronger. Thus, despite the stronger decline in aggregate activity when the ZLB binds, unconstrained firms may not be affected by the latter in equilibrium, or they may even end up producing more.

Consistently, figure 2.25 illustrates the effects of the ZLB on constrained and unconstrained firms when the economy is hit by demand, technology and financial shocks. Result are conditional on the Taylor rule considered in our analysis which proxies the way monetary policy is conducted in practice. The time preference parameter is set to 0.995 so as to imply a (lower) long-run interest rate of 2% in annualized terms (compared to 4% under baseline calibration), and the sizes of the shocks are set such that the model economy hits the ZLB. Results are reported for the UK calibration. For all three types of shocks it can be observed how the ZLB limits the positive effects of the decline in the policy rate on the production of constrained firms. It also shows how, despite the ZLB, the shock affects production of unconstrained firms either the same or even less than in its absence.The latter result is explained by the smaller negative spillover effects of monetary policy transmission to constrained firms.

Finally, one could further use the model more generally to study how optimal policy departs from the standard prescriptions derived for the credit-frictionless limit. According to the model, when ignoring the ZLB on the policy rate, such departures are only marginal in the case of the UK. Specifically, they are of order -0.08% (in annualized terms) from strict price inflation targeting for a transitory standard negative demand shock (i.e. implying an annualized variation in the efficient rate of 1%), of order 0.16% for a transitory 1% negative shock to the capital pledgeability ratio, and of order -0.025% from composite inflation targeting for a transitory positive technology shock.

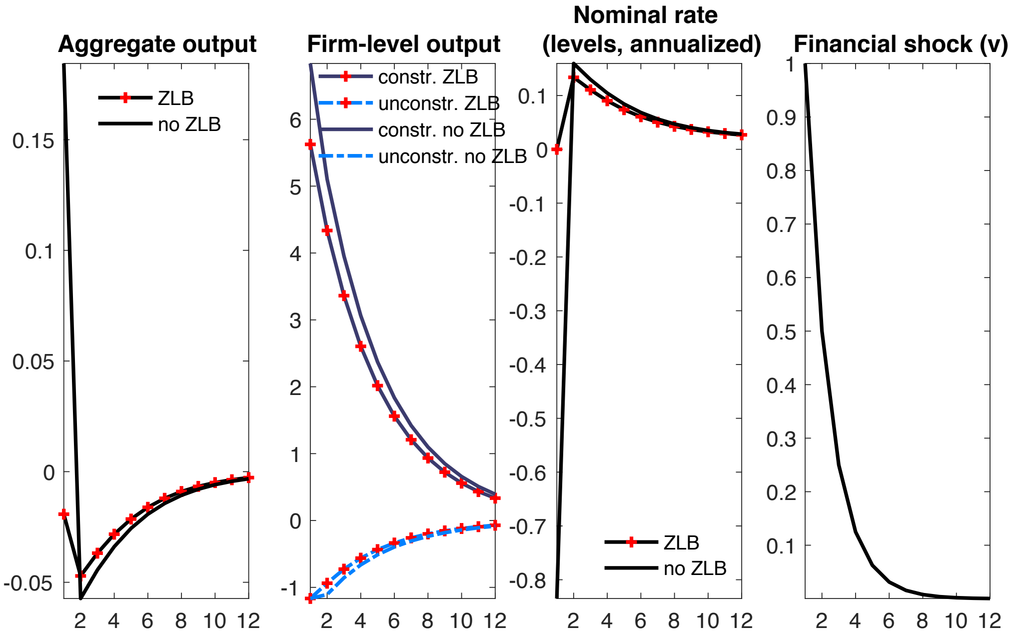


(a) Demand preference shock



(b) Technology shock

Figure 2.25: Responses to shocks subject to the ZLB



(c) Financial shock

Figure 2.25: Responses to shocks subject to the ZLB

Relatively to the credit-frictionless benchmark, the policy rate declines less in response to both demand and technology shocks under optimal policy (left and right panels in figure 2.26). Furthermore, as already pointed out in the theoretical section on optimal policy design, the policy rate declines in response to the financial shock in order to prop up collateral values (middle panel in figure 2.26).

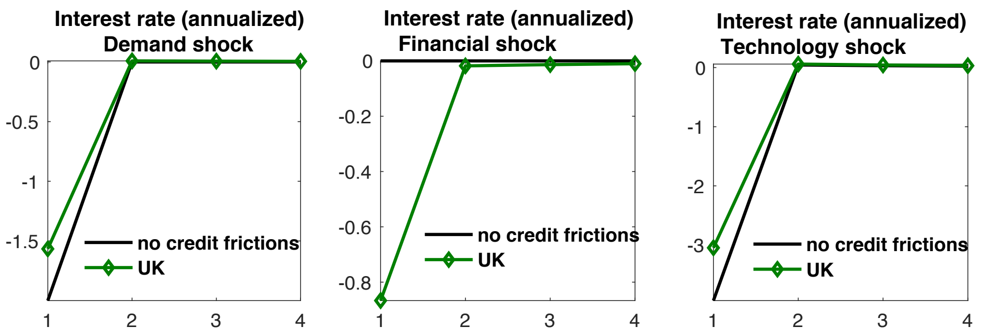


Figure 2.26: Optimal policy rate response - UK calibration and credit-frictionless setup

Going forward, since the low interest rate environment seems here to stay, it would be important to use the model to understand how monetary policy should be optimally conducted in this environment in conjunction with other available policy instruments, while taking explicitly into account the interactions between constrained and unconstrained firms in the economy.

2.9 Conclusions

In this paper I propose a stylized model to study how transmission and optimal design of monetary policy change with the share of firms facing financing constraints. The analytical framework is an extension of the Rotemberg version of the basic New Keynesian model with working capital paid in advance and physical capital in fixed aggregate supply. I find that credit frictions activate a number of additional transmission mechanisms in the New Keynesian setup through which monetary policy affects the supply-side of the economy in opposite directions. For instance, a monetary tightening depresses production by constrained firms by pushing downwards collateral values and expands it by reducing prices of inputs financed against collateral. These indirect effects spill over to unconstrained firms via input and output markets, rendering firms' heterogeneity in term of access to credit relevant for monetary policy. In equilibrium, credit frictions may both amplify or dampen the reaction of output to monetary policy depending on structural parameters, and generate "price puzzles".

Changes in transmission further translate into changes in the optimal design of monetary policy, but the latter may not be quantitatively significant unless the share of constrained firms is very high. Empirical evidence based on UK data corroborates the predictions of the model on how monetary policy affects constrained firms given the tangibility of their assets and their liquidity ratio, and uncovers a set of constrained firms whose output is steered in an unconventional direction by monetary policy. In the end, the analytical setup is used to point out that in the current low interest rate environment in countries with a dominant balance-sheet channel such as the UK (and also, the US and most likely the Euro Area), financially-constrained firms are particularly hurt, whereas unconstrained ones may benefit when the policy rate hits the ZLB.

Ongoing work focuses on bringing supporting evidence for the two new channels of monetary policy transmission identified in the theoretical analysis, namely the "input price channel" and the "spillover channel". Moreover, the empirical analysis is being extended for the US and the Euro Area.

2.10 References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2015): "House prices, collateral, and self-employment," *Journal of Financial Economics*, Vol. 117(2), 288-306.

Andrés, Javier, Oscar Arce, and Carlos Thomas (2013): "Banking competition, collateral constraints, and optimal monetary policy," *Journal of Money, Credit and Banking*, Vol. 45 (s2), 87-125.

Bahaj, Saleem, Gabor Pinter, Angus Foulis, and Paolo Surico (2019): "Employment and the collateral channel of monetary policy," Bank of England Staff Working Paper No. 827

Banerjee, Ryan, and Kristian Blickle (2018): "Financial frictions, real estate collateral, and small firm activity in Europe," FRB of New York Staff Report 868.

Bernanke, Ben S., and Mark Gertler (1995): "Inside the black box: The credit channel of monetary policy transmission," *Journal of Economic Perspectives*, 9(4), 27-48.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist (1999): "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, Vol.1, 1341-1393.

Benigno, Pierpaolo, and Michael Woodford (2012): "Linear-quadratic approximation of optimal policy problems," *Journal of Economic Theory*, Vol. 147(1), 1-42.

Bilbiie, Florin O (2019): "The new Keynesian cross," *Journal of Monetary Economics*.

Bilbiie, Florin O., Tommaso Monacelli, and Roberto Perotti (2013): "Public debt and redistribution with borrowing constraints," *The Economic Journal*, 123.566, F64-F98.

Borio, Claudio, Neale Kennedy and, Stephen D Prowse (1994): "Exploring aggregate asset price fluctuations across countries: measurement, determinants and monetary policy implications," BIS Working Papers no. 40

Calvo, Guillermo A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of monetary Economics*, 12(3), 383-398.

Carlstrom, Charles T., Timothy S. Fuerst, and Matthias Paustian (2010): "Optimal monetary policy in a model with agency costs," *Journal of Money, credit and Banking*, 42 (s1), 37-70.

Chaney, Thomas, David Sraer, and David Thesmar (2012): "The collateral channel: How real estate shocks affect corporate investment," *The American Economic Review*, 102 (6), 2381-2409.

Chodorow-Reich, Gabriel, and Antonio Falato (2017): "The loan covenant channel: How bank health transmits to the real economy," National Bureau of Economic Research No. w23879.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico (2018): "Monetary Policy, Corporate Finance and Investment," National Bureau of Economic Research No. w25366.

Debortoli, Davide, and Jordi Gali (2018): "Monetary policy with heterogeneous agents: Insights from TANK models," unpublished manuscript.

Del Negro, M. D., Eggertsson, G., Ferrero, A., and Kiyotaki, N. (2017): "The great escape? A quantitative evaluation of the Fed's liquidity facilities,"*The American Economic Review*, 107(3), 824-857.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin (2000): "Optimal monetary policy with staggered wage and price contracts," *Journal of monetary Economics*, 46 (2), 281-313.

Fort, T. C., Haltiwanger, J., Jarmin R. S., and Miranda J. (2013): "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," IMF Economic Review, 61(3), 520-559.

Faia, Ester, and Tommaso Monacelli (2007): "Optimal interest rate rules, asset prices, and credit frictions," *Journal of Economic Dynamics and control*, 31(10), 3228-3254.

Fendoglu, Salih (2014): "Optimal monetary policy rules, financial amplification, and uncertain business cycles," *Journal of Economic Dynamics and Control*, 46, 271-305.

Fiore, Fiorella De, and Oreste Tristani (2013): "Optimal monetary policy in a model of the credit channel," *The Economic Journal*, 123(571), 906-931.

Fisher, Irving (1933): "The Debt-Deflation Theory of Great Depressions", *Econometrica*, 1(4): 337.

Gali, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, second edition, Princeton University Press.

Gali, Jordi, J. David López-Salido, and Javier Vallés (2004): "Rule-of-thumb consumers and the design of interest rate rules," *Journal of Money, Credit and Banking*, 36(4), 739-764.

Gan, Jie (2007): "Collateral, debt capacity, and corporate investment: Evidence from a natural experiment," *Journal of Financial Economics*, 85(3), 709-734.

Gertler, Mark, and Simon Gilchrist (1994): "Monetary policy, business cycles, and the behavior of small manufacturing firms," *The Quarterly Journal of Economics*, 109(2), 309-340.

Gerko, Elena, and Hélene Rey (2017): "Monetary policy in the capitals of capital," *Journal of the European Economic Association*, 15(4), 721-745.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajsek (2017): "Inflation dynamics during the financial crisis," *The American Economic Review*, 107(3), 785-823.

Hansen, James (2018): ”Optimal Monetary Policy with Capital and a Financial Accelerator,” *Journal of Economic Dynamics and Control*, available online 27/04/2018.

Iacoviello, Matteo (2005): ”House prices, borrowing constraints, and monetary policy in the business cycle,” *American economic review*, 95(3), 739-764.

Jermann, Urban, and Vincenzo Quadrini (2012): ”Macroeconomic effects of financial shocks,” *The American Economic Review*, 102(1), 238-271.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor (2019): ”The effects of quasi-random monetary experiments,” *Journal of Monetary Economics*.

Kashyap, Anil K., Owen A. Lamont, and Jeremy C. Stein (1994): ”Credit conditions and the cyclical behavior of inventories,” *The Quarterly Journal of Economics*, 109(3), 565-592.

Kiyotaki, Nobuhiro, and John Moore (1997): ”Credit cycles,” *Journal of political economy*, 105(2), 211-248.

Kiyotaki, Nobuhiro (1998): ”Credit and Business Cycles”, *The Japanese Economic Review*, 49(1).

Kiyotaki, Nobuhiro, and John Moore (2002): ”Evil is the root of all money,” *The American Economic Review*, 92(2), 62-66.

Kleiner, Kristoph (2015): ”Collateral and small firm labor,” Kelley School of Business Research Paper 16-49.

Lian, Chen, and Yueran Ma (2018): ”Anatomy of corporate borrowing constraints,” Unpublished Paper, University of Chicago.

Lin, Li, Dimitrios P. Tsomocos, and Alexandros P. Vardoulakis (2015): ”Debt deflation effects of monetary policy,” *Journal of Financial Stability, Elsevier*, 21(C), 81-94.

Liu, Zheng, Pengfei Wang, and Tao Zha (2013): ”Land price dynamics and macroeconomic fluctuations,” *Econometrica*, 81(3), 1147-1184.

Lombardo, Giovanni, and David Vestin (2008): ”Welfare implications of Calvo vs. Rotemberg-pricing assumptions,” *Economics Letters*, 100(2), 275-279.

Lucas, R. E. (1981): *Models of Business Cycles*, Wiley-Blackwell.

Manea, Cristina (2018): ”Inside-money in the New Keynesian model,” unpublished manuscript

Ottonello, Pablo, and Thomas Winberry (2018): ”Financial Heterogeneity and the Investment Channel of Monetary Policy,” National Bureau of Economic Research, No. w24221.

Quadrini, Vincenzo (2011): ”Financial frictions in macroeconomic fluctuations,” *Economic Quarterly*, 97(3), 209-254.

Peters, Ryan H., and Lucian A. Taylor (2017): ”Intangible capital and the investment-q relation,” *Journal of Financial Economics* 123(2), 251-272.

Ravenna, Federico, and Carl E. Walsh (2006): ”Optimal monetary policy with the cost channel,” *Journal of Monetary Economics*, 53(2), 199-216.

Rotemberg, Julio J. (1983): "Aggregate consequences of fixed costs of price adjustment," *The American Economic Review*, 73(3), 433-436.

Schmalz, Martin C., David A. Sraer, and David Thesmar (2017): "Housing collateral and entrepreneurship," *The Journal of Finance*, 72(1), 99-132.

2.11 Appendix

2.11.1 Log-linear approximation

The analysis focuses on the first order approximation of equilibrium dynamics in the vicinity of the non-stochastic zero-inflation steady-state. Shocks are small enough for supply to be non-rationed and credit constraints to remain tight. Notation is standard: small caps stand for the log-levels, $(\widehat{\ })$ for the log-deviation from steady-state, while the absence of a time subscript denotes a steady-state value.

Households' behavior is described by the consumption/saving decision:

$$\widehat{c}_t = E_t\{\widehat{c}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\}\Big) + \frac{1}{\sigma}(1-\rho_z)z_t \tag{2.26}$$

and the (aggregate) labor supply:

$$\varphi\widehat{l}_t + \sigma\widehat{c}_t = \widehat{\omega}_t, \tag{2.27}$$

where $\widehat{\omega}_t$ is the log-deviation from steady-state of real wage. Households' behavior is not affected by the presence of credit frictions on the supply-side of the economy. Demand for goods in each set (constrained and unconstrained) equals:

$$\widehat{y}_t^s = -\varepsilon\widehat{pp}_t^s + \widehat{y}_t, \qquad s = \overline{u, c} \tag{2.28}$$

where $\widehat{pp}_t^s \equiv \widehat{p}_t^s - \widehat{p}_t$.

On the supply-side, firms' behaviour is described by their price-setting decision:

$$\pi_t^s \approx \beta E_t\{\pi_{t+1}^s\} + \lambda^s\left(\widehat{mc}_t^s - \widehat{pp}_t^s + \frac{\lambda^{1,s}}{1+\lambda^{1,s}}\widehat{\lambda}_t^{1,s}\right) \tag{2.29}$$

with $\lambda^s \equiv \frac{\varepsilon}{\xi\mathcal{M}}\frac{P^s}{P}$, $\widehat{\lambda}_t^{1,u} = 0$ and $\widehat{mc}_t^s \equiv \widehat{\omega}_t + \widehat{l}_l^s - \widehat{y}_t^s$, their production level:

$$\widehat{y}_t^s = a_t + \alpha\widehat{k}_t^s + (1-\alpha)\widehat{l}_t^s,$$

their capital demand:

$$\begin{aligned}\widehat{\varrho}_t^k \approx{}& \beta\Big(1+\lambda^{1,s}\nu\Big)E_t\{\widehat{\varrho}_{t+1}^k\} + E_t\{\widehat{\Lambda}_{t,t+1}\} - \lambda^{1,s}(1-\theta^s)\Big(\widehat{i}_t - E_t\{\pi_{t+1}\} + \widehat{\varrho}_t^k\Big)\\ &+ \beta\frac{MC(i)\big(1+\lambda^1(i)\big)\alpha Y(i)/K(i)}{Q^k/P}\Big(E_t\{\widehat{y}_{t+1}^s\} - \widehat{k}_{t+1}^s + E_t\{\widehat{mc}_{t+1}^s\}\Big) +\\ &+ \beta\lambda^{1,s}\upsilon E_t\{\widehat{\nu}_{t+1}\} + \lambda^{1,s}\left[\beta\frac{MC(i)\alpha Y(i)/K(i)}{Q^k/P} + \beta\nu - (1-\theta(i))\right]E_t\{\widehat{\lambda}_{t+1}^{1,s}\}\end{aligned}$$

where $\widehat{\varrho}_t^k \equiv log(\frac{Q_t^k}{P_t}) - log(\frac{Q^k}{P})$, $E_t\{\widehat{\Lambda}_{t,t+1}\} \approx E_t\{\Delta z_{t+1}\} - \sigma E_t\{\Delta\widehat{c}_{t+1}\}$ and, for firms in the constrained group $i \in \Theta^c$, additionally by their credit collateral constraint:

$$\widehat{\omega}_t + \widehat{l}_t^c = \widehat{k}_t^c + \frac{\nu Q^k/PK^c}{W/PL^c}\Big(\widehat{\varrho}_t^k + \nu_t\Big) - \frac{\beta^{-1}(1-\theta^c)Q^k/PK^c}{W/PL^c}\Big(\widehat{\varrho}_{t-1}^k + \widehat{i}_{t-1} - \pi_t\Big) \tag{2.30}$$

Aggregate inflation dynamics $\Pi_t \equiv \frac{P_t}{P_{t-1}}$ are approximated by:

$$\pi_t \approx \phi\bigg(\frac{P^c}{P}\bigg)^{1-\varepsilon}\pi_t^c + (1-\phi)\bigg(\frac{P^u}{P}\bigg)^{1-\varepsilon}\pi_t^u, \quad \pi_t^s \equiv log\big(\frac{P_t^s}{P_{t-1}^s}\big),\ s = \overline{u,c}, \tag{2.31}$$

whereas the goods market clearing condition by:

$$\widehat{y}_t \approx \widehat{c}_t, \qquad y_t \equiv log(Y_t) \text{ with } Y_t \equiv \Big[\int_0^1 Y_t(i)^{1-\frac{1}{\varepsilon}}di\Big]^{\frac{\varepsilon}{\varepsilon-1}}, \tag{2.32}$$

the labor market clearing condition by:

$$\widehat{l}_t \approx \phi\frac{L^c}{L}\widehat{l}_t^c + (1-\phi)\frac{L^u}{L}\widehat{l}_t^u, \tag{2.33}$$

and the one on the capital market by:

$$0 \approx \phi\frac{K^c}{K^u}\widehat{k}_{t+1}^c + (1-\phi)\widehat{k}_{t+1}^u \tag{2.34}$$

Debt market clears at the nominal interest rate set by the monetary authority(2.13).

Table 2.6: Calibration full model specification with sticky wages

Core parameters	
$\sigma = 1$	intertemporal elasticity of substitution
$\beta = 0.99$	4% steady-state real (annualized) rate
$\varphi = 5$	Frisch elasticity of labor supply 0.2
$\alpha = 0.25$	share of labor in total output 75%
$\varepsilon = 9$	steady-state price markup 12.5%
ξ	average κ equals Calvo counterpart for $\theta = 0.75$
λ_w	match the Calvo counterpart $\theta_w = 3/4$ and $\varepsilon_w = 4.5$
Credit frictions	
$\phi \in [0, 1]$	considered a variable in the analysis
Baseline	
$\alpha = 0.25$	elasticity of output to real-estate (Gali (2015))
$\nu = 0.8$	pledgeability ratio of real-estate as collateral
$\theta^c = 0.23$	fraction real-estate equity-financed by constrained firms
Alternative 1	
$\alpha = 0.03$	elasticity of output to real-estate (Iacoviello (2005))
$\nu = 0.8$	pledgeability ratio of real-estate as collateral
$\theta^c = 0.23$	fraction real-estate equity-financed by constrained firms
Alternative 2	
$\alpha = 0.03$	elasticity of output to real-estate (Iacoviello (2005))
$\nu = 0.8$	pledgeability ratio of real-estate as collateral
$\theta^c = 0.45$	fraction real-estate equity-financed by constrained firms
Shock persistence	
$\rho_m = 0$	transitory monetary impulse
$\rho_z = 0.5$	persistent demand shock
$\rho_a = 0.9$	persistent technology shock

2.11.2 The efficient allocation (following closely Gali (2015))

The efficient allocation associated with the model economy can be determined by solving the problem facing a benevolent social planner seeking to maximize the representative household's welfare, given technology and preferences. Given the absence of mechanisms for the economy as a whole to transfer resources across periods (e.g. capital accumulation), the efficient allocation corresponds to the solution of a sequence of static social planner problems. Specifically, for each period the optimal allocation must maximize the household's utility:

$$U(C_t, \{L_t(j)\}; e^{z_t}, e^{\chi_t}) \tag{2.35}$$

where $C_t \equiv \left[\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon}}di\right]^{\frac{\varepsilon}{\varepsilon-1}}$ and $L_t \equiv \int_0^1 \frac{L_t(j)^{1+\varphi}}{1+\varphi}dj$ subject to the resource constraints:

$$C_t(i) = A_t H_t(i)^\alpha L_t(i)^{1-\alpha}, \quad \forall i \in [0,1] \tag{2.36}$$

$$L_t(i) \equiv \left(\int_0^1 L_t(i,j)^{1-\frac{1}{\epsilon_w}}dj\right)^{\frac{\epsilon_w}{\epsilon_w-1}}, \quad \forall i \in [0,1] \tag{2.37}$$

$$\int_0^1 H_t(i)di = \bar{K} \tag{2.38}$$

$$\int_0^1 L_t(i)di = L_t \tag{2.39}$$

The associated optimality conditions are:

$$C_t(i) = C_t, \forall i \in [0,1] \tag{2.40}$$

$$L_t(i,j) = L_t(j) = L_t(i) = L_t, \forall i \in [0,1], \forall j \in [0,1] \tag{2.41}$$

$$H_t(i) = \bar{K}, \forall i \in [0,1] \tag{2.42}$$

$$-\frac{U_{l,t}}{U_{c,t}} = (1-\alpha)A_t\bar{K}^\alpha L_t^{1-\alpha} \tag{2.43}$$

Thus it is optimal to produce and consume the same quantity of all goods and to allocate the same amount of labor and capital to all firms. That result is a consequence of all goods entering the utility function symmetrically, combined with the concavity of utility and production functions, identical for all goods. Once the symmetric allocation is imposed, the remaining condition defining the efficient allocation, equation (2.43), equates the marginal rate of substitution between consumption and employment to the corresponding marginal rate of transformation (which in turn equals the marginal product of labor).

The NK model developed in this paper is characterized by three different distorsions: (i) the presence of market power in goods and labor markets exercised by monopolistically competitive firms and labor unions, respectively; (ii) infrequent price and wage adjustment by firms/labor unions; (iii) the presence of credit frictions affecting production of some firms.

In the decentralized economy, in the credit-frictionless limit, in the absence of nominal rigidities, (2.40), (2.41) and (2.42) are satisfied. In this case:

$$\frac{W_t}{P_t} = -\frac{U_{t,t}}{U_{c,t}}\mathcal{M}_w \tag{2.44}$$

and

$$P_t = \mathcal{M}_p \frac{W_t}{(1-\alpha)A_t\bar{K}^\alpha L_t^{1-\alpha}} \tag{2.45}$$

Note that an employment subsidy $\tau = 1 - \frac{1}{\mathcal{M}_p \mathcal{M}_w}$ funded with lump-sum taxes can be used to guaranteeing the efficiency of the flexible price/flexible wage equilibrium in the credit frictionless limit. Namely, with this subsidy in place:

$$P_t = \mathcal{M}_p \frac{(1-\tau)W_t}{(1-\alpha)A_t \bar{K}^{\alpha} L_t^{1-\alpha}} \tag{2.46}$$

and hence, condition (2.43) is also satisfied. In the welfare analysis, I assume such a subsidy is in place. Thus, the flexible-price credit-frictionless specification is the efficient benchmark at all times.

2.11.3 Zero inflation steady-state

It is convenient to first derive the steady-state shadow value of the credit collateral constraint. The model is calibrated such that the associated constraint binds, namely:

$$\left[\nu - \beta^{-1}\Big(1 - \theta^c\Big)\right]\frac{Q^k}{P}\frac{K^c}{Y^c} = (1-\tau)\frac{W}{P}\frac{L^c}{Y^c} \tag{2.47}$$

Using the price-setting equation of the constrained group of firms:

$$\frac{P^c}{P} = \mathcal{M}_p(1-\tau)\frac{W}{P}\frac{L^c}{Y^c}\frac{1+\lambda^c}{1-\alpha} \Rightarrow (1-\tau)\frac{W}{P}\frac{L^c}{Y^c} = \mathcal{M}_p^{-1}\frac{P^c}{P}\frac{1-\alpha}{1+\lambda^c} \tag{2.48}$$

and their real estate demand equation:

$$\frac{Q^k}{P} = \frac{\alpha \frac{Y^c}{K^c}\frac{P^c}{P}}{\left[\beta^{-1} - 1 - \lambda^c\Big(\nu - \beta^{-1}(1-\theta^c)\Big)\right]} \tag{2.49}$$

we can derive the equilibrium expression of λ^c exclusively as a function of structural parameters:

$$\lambda^c = \frac{(1-\alpha)\big(\beta^{-1} - 1\big)}{\mathcal{M}_p\left[\nu - \beta^{-1}\Big(1-\theta^c\Big)\right]} - \alpha \tag{2.50}$$

Next, it is convenient to compute the steady-state value $\delta^p \equiv \frac{P^c}{P^u}$. Using the production functions for the two groups of firms we obtain:

$$\frac{Y^c}{Y^u} = \Big(\frac{K^c}{K^u}\Big)^{\alpha}\Big(\frac{L^c}{L^u}\Big)^{1-\alpha} \tag{2.51}$$

Furthermore, the real estate market equilibrium $\frac{MRK^c}{P} = \frac{MRK^u}{P}$,

$$\frac{\alpha \frac{Y^u}{K^u}\frac{P^u}{P}}{\beta^{-1} - 1} = \frac{\alpha \frac{Y^c}{K^c}\frac{P^c}{P}}{\left[\beta^{-1} - 1 - \lambda^c\Big(\nu - \beta^{-1}(1-\theta^c)\Big)\right]} \tag{2.52}$$

implies:

$$\frac{K^c}{K^u} = \frac{Y^c}{Y^u}\delta^p\delta, \quad \delta \equiv \frac{\beta^{-1}-1}{\Big[\beta^{-1}-1-\lambda^c\Big(\nu-\beta^{-1}(1-\theta^c)\Big)\Big]} \tag{2.53}$$

Using $Y^u = \left(\frac{P^u}{P}\right)^{-\varepsilon} Y$ and $Y^c = \left(\frac{P^c}{P}\right)^{-\varepsilon} Y$, we can determine:

$$\frac{Y^c}{Y^u} = \left(\frac{P^c}{P^u}\right)^{-\varepsilon} = (\delta^p)^{-\varepsilon} \tag{2.54}$$

and write (2.53) as:

$$\frac{K^c}{K^u} = \Big(\delta^p\Big)^{1-\varepsilon}\delta \tag{2.55}$$

Replacing (2.55) and (2.54) in (2.51), it yields:

$$\frac{L^c}{L^u} = \Big(\delta^p\Big)^{-\frac{\varepsilon+x(1-\varepsilon)}{1-\alpha}}\Big(\delta\Big)^{-\frac{x}{1-\alpha}} \tag{2.56}$$

One way to determine δ^p is to express the tight collateral constraint (2.47) in terms of the labor-output ratio of unconstrained firms in equilibrium. The expression of L^u/Y^u as a function of L^c/Y^c can be determined by replacing the expression of K^c/K^u from (2.53) in (2.51), namely:

$$\frac{Y^c}{Y^u} = \Big(\frac{Y^c}{Y^u}(\delta^p)\delta\Big)^{\alpha}\Big(\frac{L^c}{L^u}\Big)^{1-\alpha} \Rightarrow \frac{L^c/Y^c}{L^u/Y^u} = \Big(\delta^p\delta\Big)^{-\frac{\alpha}{1-\alpha}} \tag{2.57}$$

After replacing in the binding credit collateral constraint (2.47) the expressions of L^c/Y^c from (2.57) and of $\frac{Q^k}{P}\frac{K^c}{Y^c}$ from: (2.49)

$$\frac{\Big(\nu-\beta^{-1}(1-\theta^c)\Big)\alpha\frac{P^c}{P}}{\Big[\beta^{-1}-1-\lambda^c\Big(\nu-\beta^{-1}(1-\theta^c)\Big)\Big]} = (1-\tau)\frac{W}{P}\frac{L^u}{Y^u}\Big(\delta^p\delta\Big)^{-\frac{\alpha}{1-\alpha}} \tag{2.58}$$

and using the price-setting equation of unconstrained firms:

$$\frac{P^u}{P} = \mathcal{M}_p\frac{1-\tau}{1-\alpha}\frac{W}{P}\frac{L^u}{Y^u} \Rightarrow (1-\tau)\frac{W}{P}\frac{L^u}{Y^u} = \mathcal{M}_p^{-1}(1-\alpha)\frac{P^u}{P} \tag{2.59}$$

we can compute the expression of δ^p as a function of structural parameters:

$$\delta^p = \left[\frac{1-\alpha}{\alpha}\frac{\beta^{-1}-1-\lambda^c\Big(\nu-\beta^{-1}(1-\theta^c)\Big)}{\Big(\nu-\beta^{-1}(1-\theta^c)\Big)\mathcal{M}_p}\right]^{1-\alpha}\delta^{-\alpha} \tag{2.60}$$

The value of δ^p can be directly used in (2.55) to determine $\frac{K^c}{K^u} = (\delta^p)^{1-\varepsilon}\delta$. Furthermore, using the labor market clearing condition:

$$L = \phi L^c + (1-\phi)L^u \Rightarrow 1 = \phi\frac{L^c}{L} + (1-\phi)\frac{L^u}{L} \tag{2.61}$$

and the expression of L^c as a function of L^u in (2.56), it yields:

$$\frac{L^u}{L} = \left(\phi\Big(\delta^p\Big)^{-\frac{\varepsilon+\alpha(1-\varepsilon)}{1-\alpha}}\Big(\delta\Big)^{-\frac{\alpha}{1-\alpha}} + (1-\phi)\right)^{-1} \tag{2.62}$$

and $\frac{L^c}{L} = \frac{L^u}{L}\Big(\delta^p\Big)^{-\frac{\varepsilon+\alpha(1-\varepsilon)}{1-\alpha}}\Big(\delta\Big)^{-\frac{\alpha}{1-\alpha}}$. The steady-state expression of the price index $P^{1-\varepsilon} = \phi(P^c)^{1-\varepsilon} + (1-\phi)(P^u)^{1-\varepsilon}$ implies:

$$1 = \phi\left(\frac{P^c}{P}\right)^{1-\varepsilon} + (1-\phi)\left(\frac{P^u}{P}\right)^{1-\varepsilon} \tag{2.63}$$

Using $\delta^p = \frac{P^c/P}{P^u/P}$, $\frac{P^u}{P}$ can be determined as a function of structural parameters as:

$$1 = \phi\left(\delta^p\frac{P^u}{P}\right)^{1-\varepsilon} + (1-\phi)\left(\frac{P^u}{P}\right)^{1-\varepsilon} \Rightarrow \frac{P^u}{P} = \left(\phi\Big(\delta^p\Big)^{1-\varepsilon} + (1-\phi)\right)^{\frac{1}{\varepsilon-1}} \tag{2.64}$$

and the one of $\frac{P^c}{P} = \frac{P^u}{P}\delta^p$.

In the analysis of monetary policy transmission, there is no employment subsidy correcting for market power distorsions ($\tau = 0$), whereas in the welfare analysis this subsidy is assumed to be in place.

2.11.4 Plots monetary policy transmission

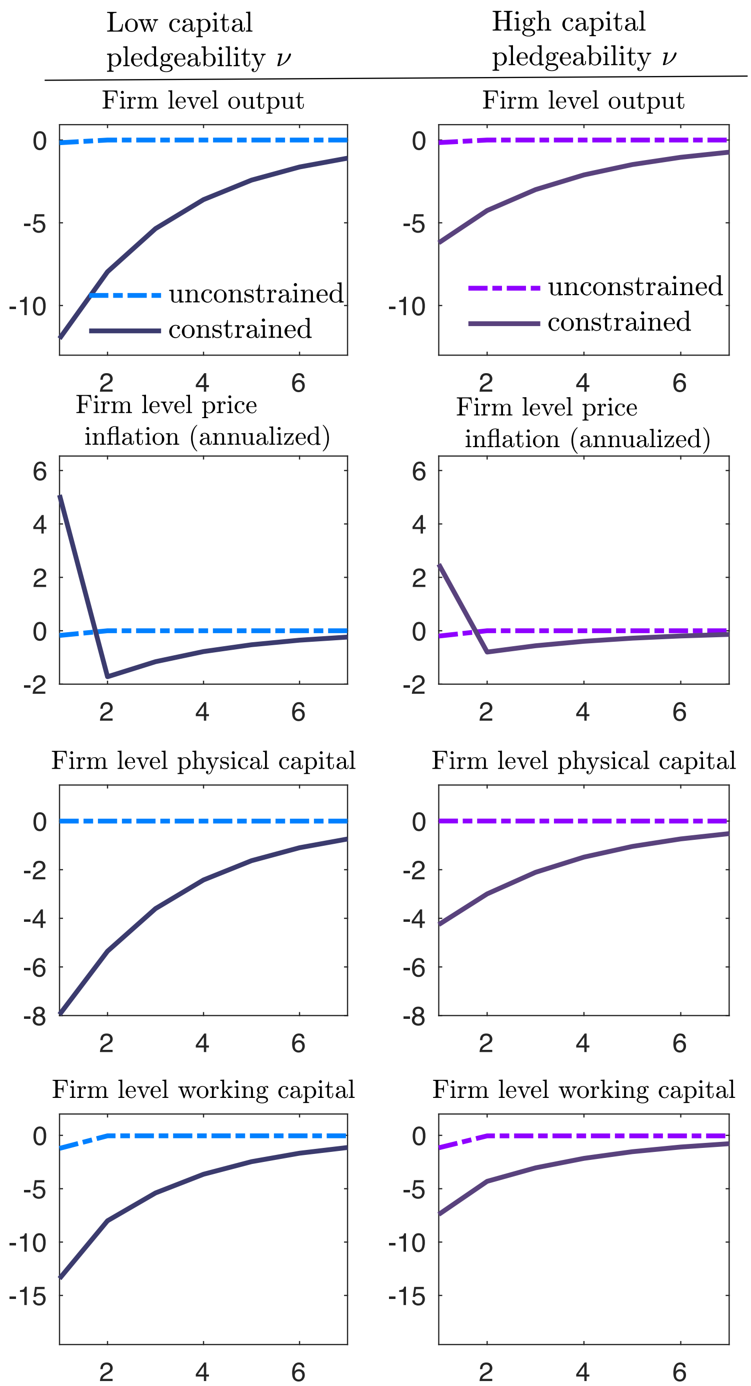


Figure 2.27: Dynamic response to a transitory monetary tightening ($\alpha = 0.03$)

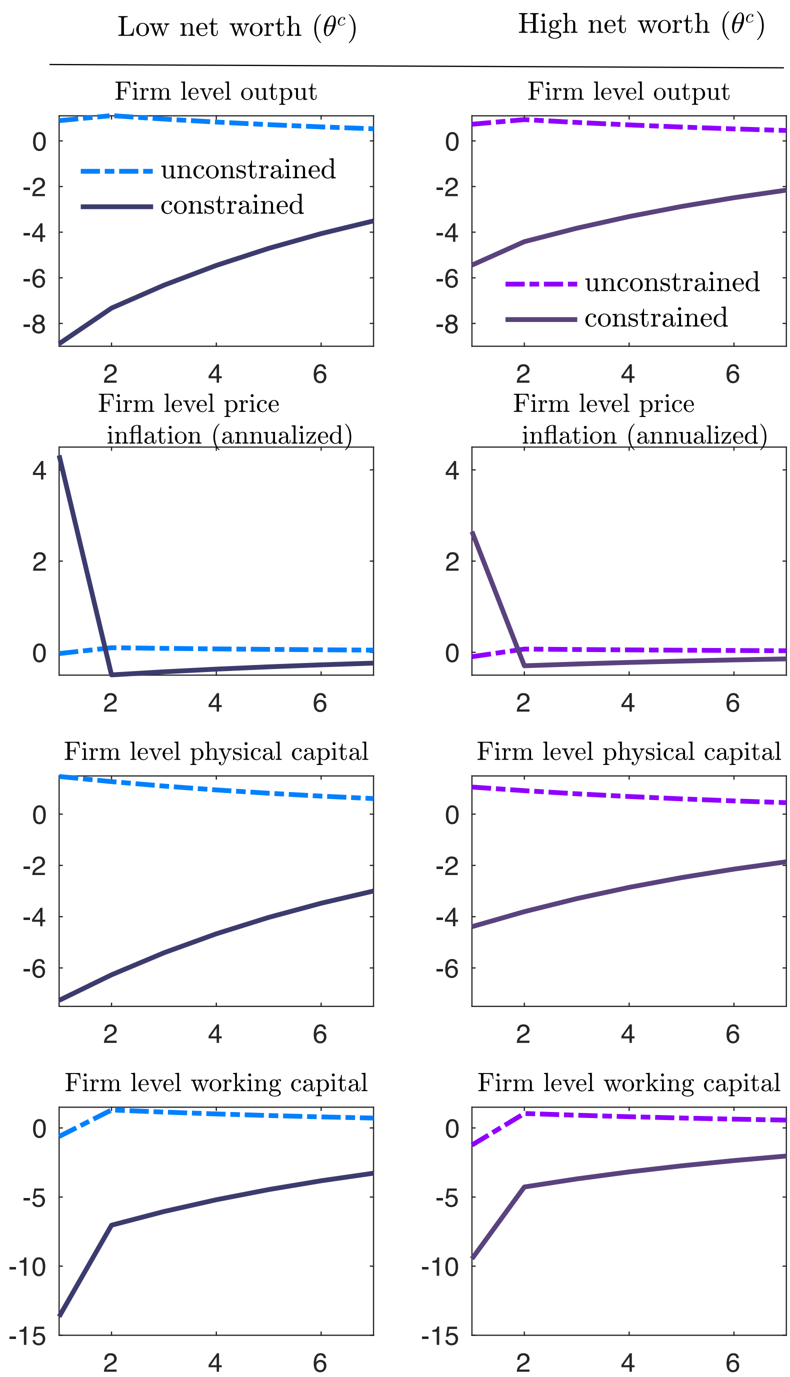


Figure 2.28: Dynamic response to a transitory monetary tightening

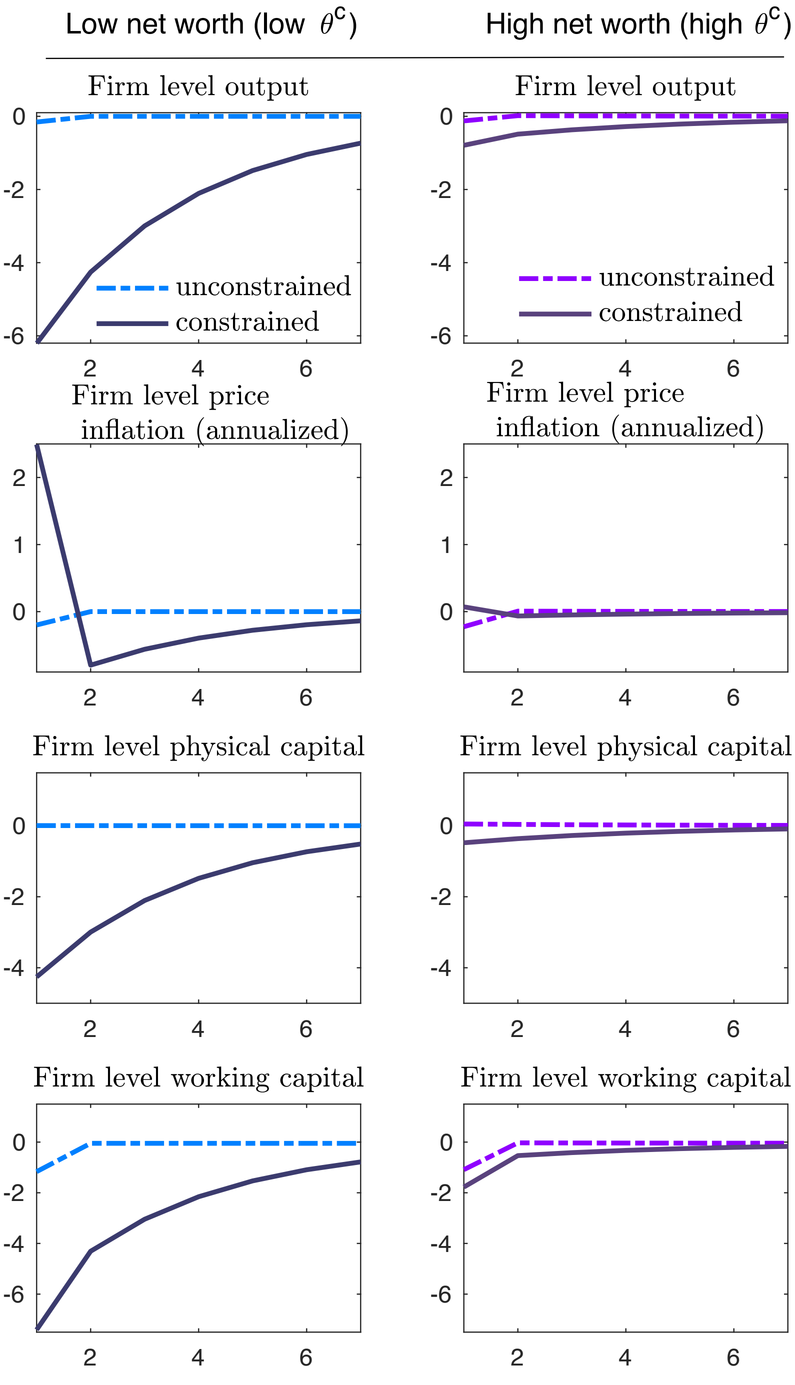


Figure 2.29: Dynamic response to a transitory monetary tightening ($\alpha = 0.03$)

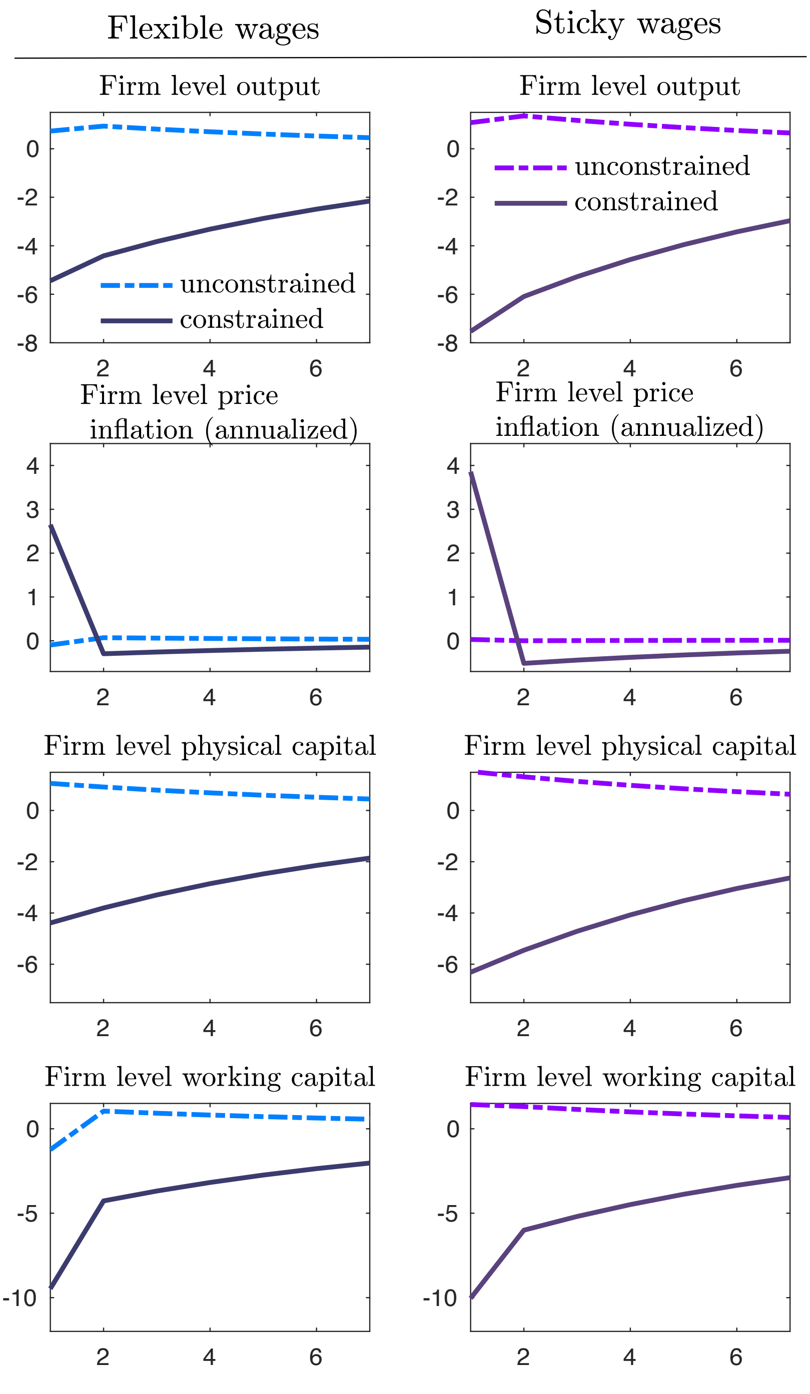


Figure 2.30: Dynamic response to a transitory monetary tightening

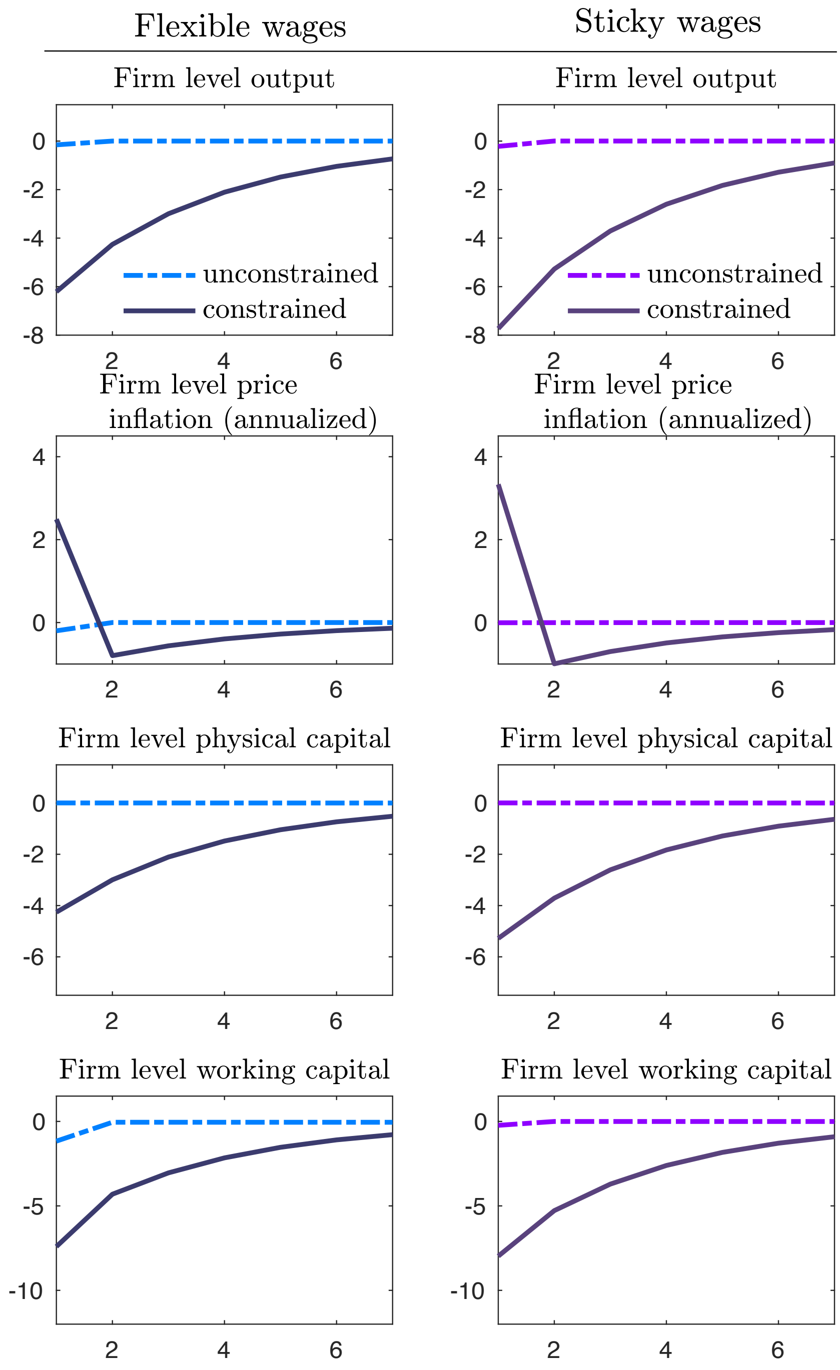


Figure 2.31: Dynamic response to a transitory monetary tightening ($\alpha = 0.03$)

2.11.5 Welfare loss function

Period utility can be written up to a second order approximation as:

$$\frac{U_t - U}{U_c Y} \approx (1+z_t)\widehat{c}_t + \frac{1-\sigma}{2}\widehat{c}_t^2 + \frac{U_l L}{U_c Y}\left((1+z_t)\int_0^1 \widehat{l}_t(j)dj + \frac{1+\varphi}{2}\int_0^1 \widehat{l}_t^2(j)dj\right) + t.i.p. \tag{2.65}$$

Next, we express $\int_0^1 \widehat{l}_t(j)dj$ and $\int_0^1 \widehat{l}_t^2(j)dj$ as a function of aggregate working hours. We define the aggregate labor hours (of all types j) used by all firms in the economy:

$$L_t \equiv \int_0^1 L_t(j)dj \tag{2.66}$$

Up to a second-order approximation, the following relations hold (Gali (2015), p. 189):

$$\int_0^1 \widehat{l}_t^2(j)dj \approx \widehat{l}_t^2 + \varepsilon_w^2 var_j\{w_t(j)\}$$

$$\int_0^1 \widehat{l}_t(j)dj \approx \widehat{l}_t - \frac{1}{2}\varepsilon_w^2 var_j\{w_t(j)\}$$

Replacing these expressions in (2.65), we obtain:

$$\begin{aligned}\frac{U_t - U}{U_c Y} \approx &(1+z_t)\widehat{c}_t + \frac{1-\sigma}{2}\widehat{c}_t^2 + \frac{U_l L}{U_c Y}\Bigg[(1+z_t)\Big(\widehat{l}_t - \frac{1}{2}\varepsilon_w^2 var_j\{w_t(j)\}\Big) + \\ &+ \frac{1+\varphi}{2}\Big(\widehat{l}_t^2 + \varepsilon_w^2 var_j\{w_t(j)\}\Big)\Bigg] + t.i.p. \\ \approx &(1+z_t)\widehat{c}_t + \frac{1-\sigma}{2}\widehat{c}_t^2 + \frac{U_l L}{U_c Y}\Bigg[(1+z_t)\widehat{l}_t + \frac{1+\varphi}{2}\widehat{l}_t^2 + \\ &+ \frac{\varphi\varepsilon_w^2}{2}\varepsilon_w^2 var_j\{w_t(j)\}\Bigg] + t.i.p.\end{aligned}$$

Next we express $\widehat{c}_t$ as a function of $\widehat{y}_t$. Up to a second order approximation the goods market clearing condition writes:

$$\widehat{c}_t \approx \widehat{y}_t - \frac{1}{2}\Big[(1-\phi)\xi(\pi_t^u)^2 + \phi\xi(\pi_t^c)^2\Big]$$

Thus,
$$\begin{aligned}\frac{U_t - U}{U_c Y} \approx &(1+z_t)\widehat{y}_t - \frac{1}{2}\Big[(1-\phi)\xi(\pi_t^u)^2 + \phi\xi(\pi_t^c)^2\Big] + \frac{1-\sigma}{2}\widehat{y}_t^2 \\ &+ \frac{U_l L}{U_c Y}\Big[(1+z_t)\widehat{l}_t + \frac{1+\varphi}{2}\widehat{l}_t^2 + \frac{\varphi\varepsilon_w^2}{2}\varepsilon_w^2 var_j\{w_t(j)\}\Big] + t.i.p.\end{aligned} \tag{2.67}$$

Next, we write $\widehat{y}_t + \frac{U_l L}{U_c Y}\widehat{l}_t$ as a function of quadratic terms[43]. L_t in (2.66) equals:

$$L_t = \int_0^1 \int_0^1 L_t(i,j)djdi = \Delta_{w,t}\Big[\phi L_t^c + (1-\phi)L_t^u\Big], \;\; \Delta_{w,t} \equiv \int_0^1 \Big(\frac{W_t(j)}{W_t}\Big)^{-\varepsilon_w} dj$$

Up to a second order approximation this relation implies:

$$\widehat{l}_t \approx \frac{\varepsilon_w}{2} var_j\{w_t(j)\} + \phi\frac{L^c}{L}\Big[\widehat{l}_t^c + \frac{1}{2}(\widehat{l}_t^c)^2\Big] + (1-\phi)\frac{L^u}{L}\Big[\widehat{l}_t^u + \frac{1}{2}(\widehat{l}_t^u)^2\Big]$$

where $var_j\{w_t(j)\}$ defines the second order approximation of $\int_0^1 (W_t(j) - W_t)^2 dj$. Using the expression of aggregate output, the production functions, and the expression above, $\widehat{y}_t + \frac{U_l L}{U_c Y}\widehat{l}_t$ equals:

$$\begin{aligned}\widehat{y}_t + \frac{U_l L}{U_c Y}\widehat{l}_t =& \phi\Big(\frac{P^c}{P}\Big)^{1-\varepsilon}\widehat{y}_t^c + (1-\phi)\Big(\frac{P^u}{P}\Big)^{1-\varepsilon}\widehat{y}_t^u + \\ &+ \frac{U_l L}{U_c Y}\Bigg[\frac{\varepsilon_w}{2} var_j\{w_t(j)\} + \phi\frac{1}{2}\frac{L^c}{L}(\widehat{l}_t^c)^2 + \frac{1}{2}(1-\phi)\frac{L^u}{L}(\widehat{l}_t^u)^2 + \\ &+ \phi\frac{L^c}{L}\frac{\widehat{y}_t^c - a_t - \alpha\widehat{k}_t^c}{1-\alpha} + (1-\phi)\frac{L^u}{L}\frac{\widehat{y}_t^u - a_t - \alpha\widehat{k}_t^u}{1-\alpha}\Bigg] + t.i.p.\end{aligned}$$

[43] I tried first to express $\widehat{l}_t$ as a function of aggregate output. In the basic credit-frictionless NK model $L_t = \Delta_{w,t}\Delta_{p,t}\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}$ with $\Delta_{w,t} \equiv \int_0^1 \left(\frac{W_t(j)}{W_t}\right)^{-\varepsilon_w} dj$, $\Delta_{p,t} \equiv \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{\frac{-\varepsilon_p}{1-\alpha}} di$. Hence, up to a second order approximation, $(1-\alpha)\widehat{l}_t = \widehat{y}_t - a_t + \frac{(1-\alpha)\varepsilon_w}{2} varj\{w_t(j)\} + \frac{\varepsilon_p}{2\Theta} vari\{p_t(i)\}$ (Gali (2015), Chapter 6, page 190). Here, due to the presence of capital in the production function, and of the different capital levels used by unconstrained versus constrained firms, $L_t = \Delta_{w,t}\Delta_{pk,t}\left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}$, with $\Delta_{pk,t} \equiv \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{\frac{-\varepsilon_p}{1-\alpha}} K_t(i)^{-\frac{\alpha}{1-\alpha}} di$. Thus, in the second order approximation of the expression of aggregate labor there are first and second order terms including $\widehat{k}_t^u$ and $\widehat{k}_t^c$. The approach followed is less computational intensive.

Replacing the expression above in (2.67), we get (after some algebra),

$$\begin{aligned}
\frac{U_t - U}{U_c Y} \approx & -\frac{1}{2}\Big[\xi(1-\phi)(\pi_t^u)^2 + \xi\phi(\pi_t^c)^2 - (1-\sigma)\widehat{y}_t^2 + \\
& + \frac{(1-\alpha)(1+\varphi)}{L^u/L}\left(\frac{P^u}{P}\right)^{1-e}\widehat{l}_t^2 + \\
& + \frac{(1-\alpha)e_w(1+\varphi e_w)}{L^u/L}\left(\frac{P^u}{P}\right)^{1-e} var_j\{w_t(j)\} + \\
& + \frac{(1-\alpha)}{L^u/L}\left(\frac{P^u}{P}\right)^{1-\varepsilon}\Big[\phi\frac{L^c}{L}(\widehat{l}_t^c)^2 + (1-\phi)\frac{L^u}{L}(\widehat{l}_t^u)^2\Big] - \\
& - 2\phi(\frac{P^c}{P})^{1-\epsilon}\frac{\lambda^c}{1+\lambda^c}\widehat{y}_t^c - 2\alpha\phi\left((\frac{P^c}{P})^{1-\epsilon}\frac{1}{1+\lambda^c} - (\frac{P^u}{P})^{1-\epsilon}\frac{K^c}{K^u}\right)\widehat{k}_t^c \\
& + \alpha(\frac{P^u}{P})^{1-\epsilon}\left(\frac{K^c}{K^u}\phi(\widehat{k}_t^c)^2 + (1-\phi)(\widehat{k}_t^u)^2\right)\Big] + t.i.p.
\end{aligned}$$

We can thus define the loss function as:

$$\begin{aligned}
\mathcal{L}_0 \approx & - E_0\sum_{t=0}^{\infty}\beta^t\frac{U_t - U}{U_c Y} \\
= & \frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t\Big[\xi(1-\phi)(\pi_t^u)^2 + \xi\phi(\pi_t^c)^2 - (1-\sigma)\widehat{y}_t^2 + \gamma^l\widehat{l}_t^2 + \gamma^w(\widehat{\pi}_t^w)^2 + \\
& + \gamma^{lcu}\left(\phi\frac{L^c}{L}(\widehat{l}_t^c)^2 + (1-\phi)\frac{L^u}{L}(\widehat{l}_t^u)^2\right) - \gamma^{yc}\widehat{y}_t^c - \gamma^{kc}\widehat{k}_t^c + \\
& + \gamma^{kcu}\left(\frac{K^c}{K^u}\phi(\widehat{k}_t^c)^2 + (1-\phi)(\widehat{k}_t^u)^2\right)\Big] + t.i.p.
\end{aligned}$$

where I used $\sum_{t=0}^{\infty}\beta^t var_j\{w_t(j)\} = \frac{\theta^w}{(1-\beta\theta^w)(1-\theta^w)}(\pi_t^w)^2$ (see Galí (2015), page 119): $\gamma^l \equiv \frac{(1-\alpha)(1+\varphi)}{L^u/L}(\frac{P^u}{P})^{1-\epsilon} \geq 0$,
$\gamma^w \equiv \frac{(1-\alpha)\epsilon^w(1+\varphi\epsilon^w)\theta^w}{L^u/L(1-\beta\theta^w)(1-\theta^w)}(\frac{P^u}{P})^{1-\epsilon} \geq 0$,
$\gamma^{lcu} \equiv \frac{(1-\alpha)}{L^u/L}(\frac{P^u}{P})^{1-\epsilon} \geq 0$,
$\gamma^{yc} \equiv 2\phi(\frac{P^c}{P})^{1-\epsilon}\frac{\lambda^c}{1+\lambda^c} \geq 0$,
$\gamma^{kc} \equiv 2\alpha\phi\left((\frac{P^c}{P})^{1-\epsilon}\frac{1}{1+\lambda^c} - (\frac{P^u}{P})^{1-\epsilon}\frac{K^c}{K^u}\right) \geq 0$,
$\gamma^{kcu} \equiv \alpha(\frac{P^u}{P})^{1-\epsilon} \geq 0$,
$t.i.p.$ collects terms that are independent of monetary policy. I derive the optimal monetary policy under commitment by choosing all variables in section to minimize the welfare loss function above subject to the equations of the model in section 2.11.1.

Table 2.7: Coefficients linear terms welfare criterion

	$\phi = 0$	$\phi = 0.2$	$\phi = 0.5$
Baseline			
γ^{yc}	0	0.017	0.046
γ^{kc}	0	-0.015	-0.041
Alternative 2			
γ^{yc}	0	0.0019	0.0051
γ^{kc}	0	-0.0017	-0.0045

2.11.6 Plots optimal monetary policy

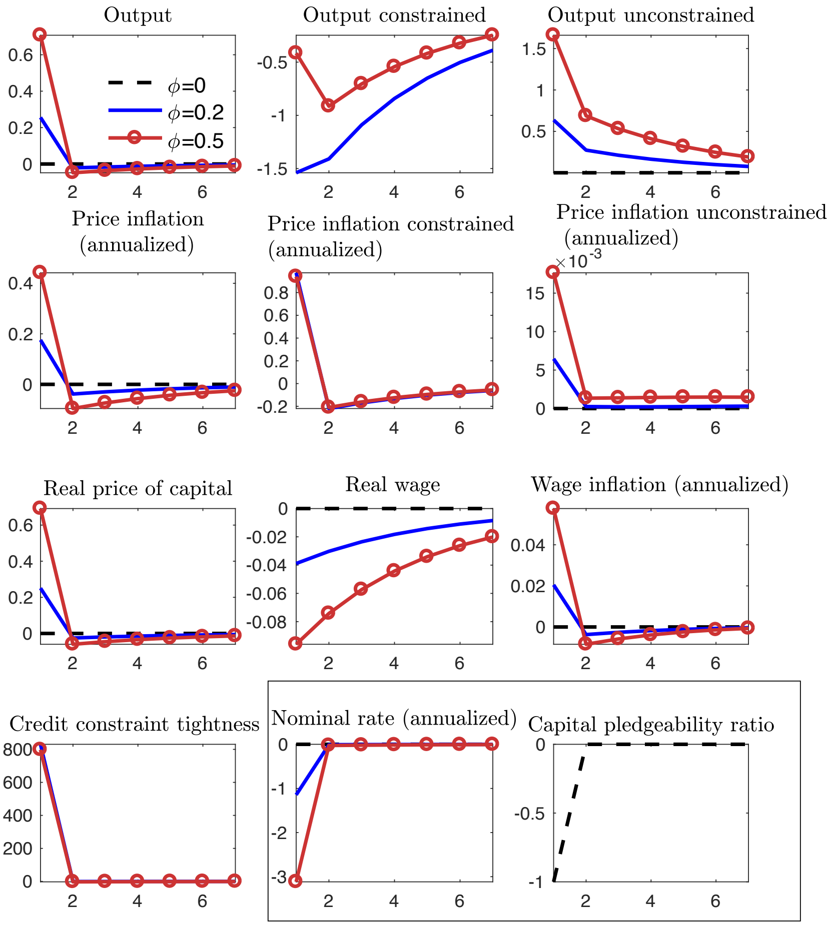


Figure 2.32: Optimal policy: Dynamic responses to a transitory negative pledgeability ratio shock for different shares of constrained firms ($\alpha = 0.03$, $\theta^c = 0.45$)

Chapter 3

OPTIMAL MONETARY-FISCAL POLICY WITH FISCAL LIMIT AND ZERO LOWER BOUND CONSTRAINTS

3.1 Abstract

With their aging populations, many advanced economies are currently (i) approaching their fiscal limit, and (ii) facing a decline in their long-run real interest rates. The latter increases the probability that the policy rate is constrained by the zero lower bound (ZLB). In this paper, we study the optimal monetary-fiscal policy mix under commitment in the presence of fiscal limit and ZLB constraints. We conduct our analysis in an extension of the basic New Keynesian model with an endogenous fiscal limit. We assume away both outright default on public debt and monetary financing. Our main result is that, as the economy approaches its fiscal limit, the reduction in fiscal space limits the future boom that the policy-maker can promise at the ZLB, and hence dynamics under optimal policy become less inflationary[1].

Keywords: Monetary policy, fiscal policy, endogenous fiscal limit, zero lower bound **JEL Class.:** E2 – E3 – E4.

[1]Joint work with Alain Durré, Université Paris Dauphine and Université Catholique de Lille. We thank Jordi Galí for guidance and encouragement to pursue this project, as well as for useful comments to Andrea Caggese, Alberto Martin, Edouard Schaal and participants at the CREI Macro Lunch Seminar in November 2018.

3.2 Introduction

With their aging populations, many advanced economies are currently (i) approaching their fiscal limit, and (ii) facing a decline in their long-run real interest rates. The latter increases the probability that the policy rate is constrained by the zero lower bound (ZLB). The aging of populations have been put forward as a common driving force for both long-run phenomena by two separate strands of literature. One is the *fiscal limit* literature which warns that the rise in old-age dependency ratios in advanced market economies (figure 3.1) requires high tax adjustments to finance adequate pensions, health and long-term care expenditures (figure 3.2), and thus pushes government finances on unsustainable paths (e.g. Davig, Leeper and Walker (2010), Leeper and Walker (2011)). The other is the *secular stagnation* literature (e.g. Summers (2014), Eggertsson and Mehrotra (2014), Gordon (2016)) which outlines that these demographic changes, alongside other factors, push real long-term interest rates downwards (figure 3.3). This is because the decrease in the number of young relative to middle aged reduces loan demand, and hence triggers a decline in real interest rates (Eggertsson, Mehrotra and Robbins (2019)).

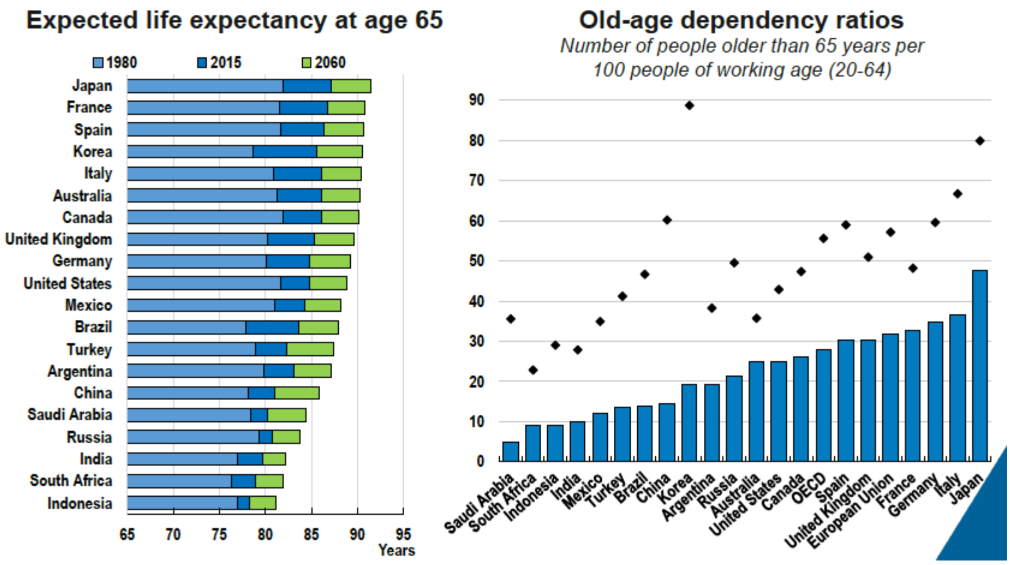


Figure 3.1: Aging populations in OECD countries (◇ projected values for 2060)
Source: OECD (2019) – Fiscal changes and inclusive growth in aging societies

Our paper explores how the two ongoing long-run phenomena (the approach to the fiscal limit and the decline in long-run interest rates) jointly affect the optimal monetary-fiscal response to business cycle fluctuations. To investigate our research question, we use an extension of the basic New Keynesian (NK) model (Galí (2015), Woodford (2003)) with an endogenous fiscal limit and we study how the optimal policy mix at the ZLB changes with the proximity of the economy to its long-run fiscal limit. We assume away both outright default on public debt and monetary financing. Our main result is that as the economy approaches its fiscal limit, the reduction in fiscal space limits the future boom that the policymaker can promise at the ZLB, and hence dynamics under optimal policy become less inflationary. We also find that, on this convergence path, it gradually becomes more costly for the policymaker to correct for long-run allocation inefficiencies associated to the increase in distorsionary taxes. And this even in the absence of the ZLB.

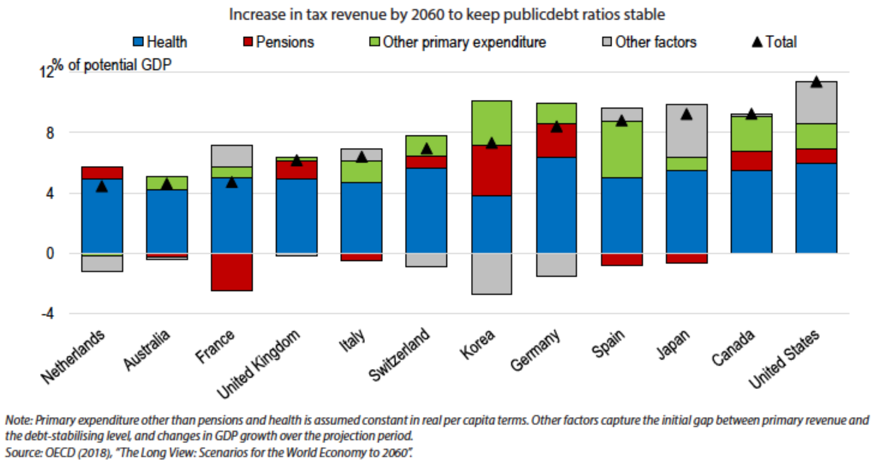


Figure 3.2: Steep tax increases needed to absorb ageing costs
Source: OECD (2019) – Fiscal changes and inclusive growth in aging societies

So far, literature has investigated separately how the two long-run phenomena affect the conduct of monetary and fiscal policies. Specifically, on the one hand, the "fiscal limit" literature has focused on the implications of the long-run unsustainability of public finances for monetary policy without reference to business cycle stabilization and the ZLB (e.g. Bianchi and Melosi (2018), Leeper (2013), Davig, Leeper, and Walker (2011), Leeper and Walker (2011))[2]. In a nutshell,

[2]Leeper (2013) assumes the economy is already at the fiscal limit where agents expect it to remain forever and studies how the economy operates when the policymaker opts to allow inflation to devalue outstanding government debt. Davig, Leeper, and Walker (2011) and Leeper and Walker (2011) incorporate endogenous fiscal limits triggered by Laffer-curve relationships as a result of distorsionary taxation and/or exogenous fiscal limit distributions triggered by political factors given (current and future) public transfers/expenditures commitments and study how

these papers argue that the monetary authority will need to embark on a more inflationary regime so as to render (nominal) debt sustainable by eroding its real value. On the other hand, Eggertsson and Woodford (2004) and Nakata (2017) have derived the optimal monetary-fiscal policy mix when the ZLB binds without taking the fiscal limit into account[3]. They find that fiscal and monetary policies should be jointly used to promise a future boom at the ZLB. This future boom is aimed to encourage present consumption, and hence to compensate for the lack of ammunition of monetary policy when the shock arises.

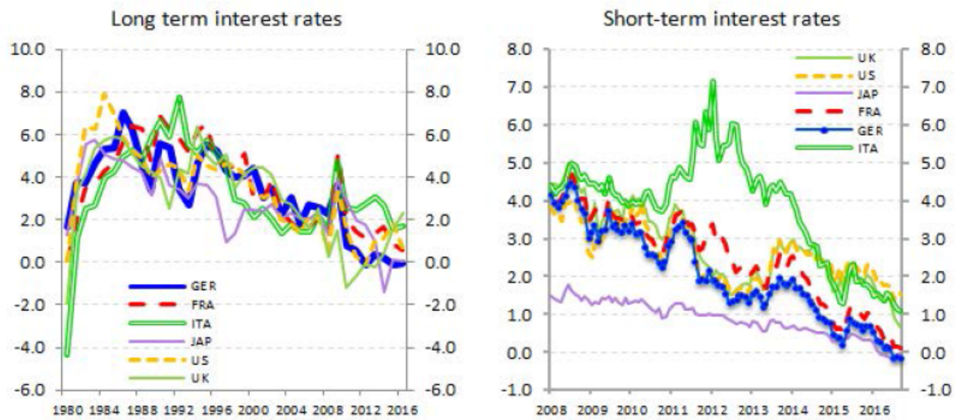


Figure 3.3: Real interest rates in main advanced economies
Source: Ferrero, Gross and Neri (2019)

Our analysis which considers jointly the policy implications of the two ongoing long-run phenomena uncovers a number of novel insights. On the one hand, relatively to the ZLB literature, we find that as the economy approaches its fiscal limit, the magnitude of the future boom that the policymaker can promise at the ZLB under the optimal monetary-fiscal policy mix decreases[4]. On the other

possible regime switches at the fiscal limit may affect current macroeconomic behaviour. Leeper and Leith (2016) look at the optimal monetary-fiscal response to transfer and public expenditure shocks for different levels of initial public debt; their analysis however does not take into account neither standard business-cycle shocks (e.g. demand or technology ones), nor the zero lower bound constraint on the policy rate.

[3]This latter reference builds on previous results from the literature on optimal monetary-fiscal policy mix which abstracts from the existence of a ZLB (Benigno and Woodford (2003), Schmitt-Grohe and Uribe (2004), Siu (2004)).

[4]As in Eggertsson and Woodford (2004) and Nakata (2017), we implicitly assume that structural parameters of the model and the shocks are such that a zero-inflation steady-state is optimal. From the analyses in Budianto (2019) and Andrade et al. (2018), however, we can see that this is not always the case.

hand, relatively to the 'fiscal limit' literature, we find that fiscal pressures can affect monetary policy even when they don't endanger the sustainability of public finances. In particular, when the ZLB binds, these fiscal pressures restrict the stabilization power of the monetary-fiscal policy mix at business cycle frequency[5].

In building our analytical setup, we model explicitly the two on-going low-frequency phenomena triggered by demographic changes as affecting the long-run equilibrium of the model. In particular, a lower long-run interest rate is mapped to a lower value of the steady-state interest rate as in the ZLB literature (pinned down by a lower time discount factor), whereas the relevant fiscal limit in the analysis is the steady-state fiscal limit. To add an endogenous fiscal limit into the basic NK model (Galí (2015), Woodford (2003)) we embed a "Laffer curve" by assuming that taxes are distorsionary instead of lump-sum, whereas to gauge the proximity of the economy to its long-run fiscal limit we add exogenous (real) steady-state transfers. The latter stand for pre-committed expenses associated to aging populations such as pensions and healthcare expenditures. In modeling the fiscal pressures due to aging populations (depicted in figure 3.3) as higher transfers to households, we follow the fiscal-limit literature (e.g. Bi (2012), Davig, Leeper, and Walker (2010)). In our setup, higher steady-state transfers bring closer the economy to its steady-state (long-run) fiscal limit, and may push it beyond. Apart from these two features, the model is similar to the basic NK model on all other dimensions.

We start our analysis by looking how the distance of the economy to its long-run fiscal limit (defined as the peak of the Laffer curve in steady-state) affects optimal policy *in the absence of the ZLB constraint*. We first point out that, as the economy converges to its fiscal limit, the higher required levels of distorsionary taxation widen long-run allocation inefficiencies. Abstracting from exogenous disturbances, we find that monetary and fiscal policies can (partially) correct these distorsions, and to do so, they need to coordinate. If prices were fully-flexible, (long-run) efficiency could be restored under optimal policy by using unexpected inflation to permanently reduce the real value of (nominal) public debt, and hence debt service costs. This would make room for a permanent decline in tax rates which would restore the equality between output and its efficient level. When prices are sticky, optimal policy still lowers the value of long-run debt so as to allow for lower tax rates and higher output in the long run. The way the permanent decline in long-run debt is now achieved depends however on the degree of price stickiness.

Particularly, when price stickiness is low, the permanent decline in real debt

[5]Moreover, unrelated to the ZLB, but related to optimal business cycle stabilization policies, we also point out that (i) current fiscal pressures also widen allocation inefficiencies as they imply levying greater amounts of distortionary taxes, and that, (ii) as the economy approaches its fiscal limit, correcting such inefficiencies becomes more costly.

is still achieved via unexpected inflation as in the case of fully-flexible prices. However, the policymaker can no longer necessarily restore efficiency because (i) it is constrained to choose an inflation which is compatible not only with the government's budget constraint, but also with firms' price-setting behaviour, and (ii) inflation has welfare costs. Furthermore, for standard degrees of price stickiness, the policymaker engineers instead the initial decline in real debt via both unexpected inflation and higher real taxes. To achieve this, it initially simultaneously declines the policy rate and increases the tax rate. The decline in the policy rate temporary increases aggregate demand and hence counteracts the initial negative effects of the tax hike on output. This makes possible a simultaneous initial increase in real taxes and output (which would not be possible if prices were flexible). Alongside the rise in real tax income, unexpected inflation contributes as well to the decline in real debt, albeit to a significantly lower extent than in the case of flexible prices. The latter arises as a result of higher demand (due to the decline in the policy rate), and lower supply (due to the tax hike). As the economy approaches its fiscal limit, the ability to engineer the initial *increase* in taxes via the increase in the tax rate (with its associated debt reduction) diminishes. Using inflation remains an option, but this makes more costly to correct for long-run allocation inefficiencies.

In a next step, we study the optimal policy response to a demand preference shock. We choose this particular type of shock because we impose the ZLB constraint later in our analysis, and this shock is specific to the ZLB literature. We find that monetary and fiscal policies coordinate again to engineer the optimal response and that, while doing so, they take into account as well the need to correct for long-run distorsions. Specifically, the policy rate is optimally used to engineer a monetary policy stance slightly looser than the one warranted to perfectly offset the effect of the shock on aggregate demand. The more looser policy pushes (directly) output above its initial steady-state level by expanding aggregate demand and erodes the real value of debt permanently via inflation. Simultaneously, fiscal policy varies the tax rate such that, given the equilibrium effects induced by the variation in the policy rate, long-run output increases by the optimal amount. The variation in the tax rate depends on the sign of the shock: for a negative shock, the decline in the policy rate has an initial direct negative effect on debt and hence there is less need to raise taxes on impact (and, hence, to increase the tax rate). The converse is true for a positive shock when the rise in the policy rate puts additional positive pressures on debt.

Once we have understood how the fiscal limit affects optimal policy in the absence of the ZLB, we turn to our main research question and study how conclusions change when we also add this constraint on the policy rate. In this case, we find that the labor income tax is additionally actively used (alongside the policy rate) to counteract the effects of negative demand shocks. Specifically, as in

Eggertsson and Woodford (2004) and Nakata (2016), the policymaker initially increases the tax rate to act against deflationary pressures, and it then commits to keep it low even in the aftermath of the shock so to engineer a future boom, side by side with the promised accommodative future monetary policy. The 'promissed' future boom engineered via both loose monetary and fiscal policies is aimed to encourage present consumption. The fiscal limit constrains the extent to which tax policy can be used as a substitute for monetary policy at the ZLB. Specifically, it limits the rise in tax income on impact and, hence, the level of accommodative policy that the policymaker can promise for the future. This is because, for each marginal variation in the tax rate when the shock arises, when the economy is closer to its (long-run) fiscal limit, taxes increase less or even decline. As a result, for each initial marginal increase in taxes, there is less fiscal space created that can be used to engineer a "promised" boom in later periods without endangering the sustainability of public finances. This is why, when the ZLB binds, dynamics under optimal policy mix in response to a negative demand shock are less inflationary as the economy approaches the fiscal limit[6].

Hereafter, the paper is organized as follows: section 2 describes the model, section 3 derives dynamics under the optimal policy mix– first without the ZLB constraint (subsection 3.1) and then with this constraint (subsection 3.2), whereas section 4 concludes by summarizing main results and discussing future extensions.

3.3 Model

The model is an extension of the cashless-limit basic New Keynesian (NK) setup (e.g. Chapters 3 in Galí (2015) and Woodford (2003)) with a (distorsionary) labor income tax and exogenous steady-state (real) lump-sum transfers to households. The two features generate an endogenous Laffer curve in equilibrium, and the latter allows to gauge the distance of the economy to its long-run fiscal limit (defined as the peak of the Laffer curve). The agents in the model are a continuum of size one of identical households, a continuum of size one of monopolistic firms, and a public authority setting jointly both monetary and fiscal policies.

3.3.1 Households

The representative household decides each period how much to consume C_t, to work L_t and to invest in one-period nominal public bonds B_t in order to maximize

[6]Note that as the policy mix affects the real value of nominal debt. Thus, default via the erosion of the real value of nominal debt is allowed.

expected inter-temporal lifetime utility

$$E_0\left\{\sum_{t=0}^{\infty}\beta^t Z_t\left(\frac{C_t^{1-\sigma}}{1-\sigma}-\frac{L_t^{1+\varphi}}{1+\varphi}\right)\right\}$$

subject to the budget constraint

$$P_tC_t+Q_tB_t\leq B_{t-1}+(1-\tau_t)W_tL_t+Div_t+P_t\mathcal{T}\ \ \forall t, \tag{3.1}$$

and the solvency (transversality) condition

$$\lim_{T\to\infty}E_0\left\{\beta^{T-t}\frac{Z_TC_T^{-\sigma}B_T}{Z_tC_t^{-\sigma}P_T}\right\}\geq 0 \tag{3.2}$$

where $C_t\equiv[\int_0^1 C_t(i)^{1-\frac{1}{\varepsilon}}di]^{\frac{\varepsilon}{\varepsilon-1}}$ is a standard Dixit-Stiglitz consumption index of differentiated goods with ε a measure of substitutability among them, $P_t\equiv[\int_0^1 P_t(i)^{1-\varepsilon}di]^{\frac{1}{1-\varepsilon}}$ is the unit price of the consumption basket, W_t stands for the nominal wage, B_t represents purchases of one-period government bonds sold at price Q_t which pay one nominal unit at $t+1$, τ_t is a (time-varying) tax rate on labor income, Div_t are dividends from the ownership of firms, $\mathcal{T}$ are real transfers from the public sector, and $log(Z_t)=\rho_z log(Z_{t-1})+\varepsilon_t^z$ is a demand preference shock. The optimality conditions describing the behaviour of the household are the sequence of consumption/saving decisions,

$$Q_t=\beta E_t\left\{\frac{C_{t+1}^{-\sigma}Z_{t+1}}{C_t^{-\sigma}Z_t}\frac{P_t}{P_{t+1}}\right\}\ \ \forall t, \tag{3.3}$$

labor supply ones

$$C_t^{\sigma}L_t^{\varphi}=(1-\tau_t)\frac{W_t}{P_t}\ \ \forall t, \tag{3.4}$$

together with the solvency condition (3.2), and the period budget constraints (3.1). Up to a first order log-linear approximation, these conditions write

$$\widehat{c}_t=E_t\{\widehat{c}_{t+1}\}-\frac{1}{\sigma}(\widehat{i}_t-E_t\{\pi_{t+1}\})+\frac{1}{\sigma}(1-\rho_z)z_t,\quad\forall t, \tag{3.5}$$

$$\widehat{\omega}_t=\sigma\widehat{c}_t+\varphi\widehat{l}_t+\frac{\tau}{1-\tau}\widehat{\tau}_t,\quad\forall t, \tag{3.6}$$

where $\widehat{i}_t\equiv -log(Q_t)-(-log(Q))$ is the deviation from steady-state of the nominal bond yield, $\omega_t\equiv log(\frac{W_t}{P_t})-log(\frac{W}{P})$ is the log of the real wage rate. The representative household never borrows in equilibrium, and hence its solvency requirement (3.2) is always satisfied. Same for its budget constraints (3.1) given “Walras-Law”.

3.3.2 Firms

Firms are in monopolistic competition and each of them produces a different variety i. They are all endowed with an identical Cobb-Douglas production technology

$$Y_t(i) = AL_t^{1-\alpha}(i), \tag{3.7}$$

and face Calvo-type price adjustment constraints. In this environment, at each date only θ of them reset their prices according to the optimality condition

$$p_t^* = \mu + (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k E_t\{\psi_{t+k/t}\} \quad \forall t, \tag{3.8}$$

where p_t^* is the log of the optimal price, μ is the log of the desired gross markup $\mathcal{M} \equiv \frac{\epsilon}{\epsilon-1}$, $\psi_{t+k/t} \equiv -log(1-\alpha) + (\omega_t + p_t) + l_{t+k/t} - y_{t+k/t}$ is the log of the nominal marginal cost at date $t+k$ of a firm which last reset its price at date t (in logs). The remaining fraction $1-\theta$ of firms satiates its demand at the posted price. A detailed derivation of the price setting equation can be found for instance in Chapter 3 in Galí (2015) or Woodford (2003).

3.3.3 Public sector

A benevolent monetary-fiscal authority choses optimally the labor income tax rate τ_t, the volume of public debt B_t and its price Q_t, given pre-committed (exogenous) real period (lump-sum) transfers $\mathcal{T}$. Policy variables satisfy at each date t the flow budget constraint

$$Q_t B_t = B_{t-1} + (P_t\mathcal{T} - \tau_t W_t L_t), \tag{3.9}$$

and the solvency condition

$$\lim_{T\to\infty} \beta^T E_t\{Q_T B_T\} = 0 \tag{3.10}$$

The flow budget-constraint constraint states that newly issued public debt at each date equals outstanding debt plus the current fiscal deficit. The solvency constraint states that government debt cannot increase at a rate faster than the interest rate. The government can credibly promise to repay its (nominal) debt every period and can also commit to future policy actions. Also, it stands ready to issue non-interest bearing cash-balances, implying that the highest price it can set on the nominal bond is one, hence $Q_t \leq 1$.

Hereafter, we will use in the analysis the log-linearized versions of the two constraints around the non-stochastic zero-inflation steady-state[7]

$$\widehat{b}_t = \beta^{-1}\Big(\widehat{b}_{t-1} - \pi_t - \frac{\tau W/PL}{B/P}(\widehat{\tau}_t + \widehat{\omega}_t + \widehat{l}_t)\Big) + \widehat{i}_t \quad \forall t, \tag{3.11}$$

$$\lim_{T\to\infty} \beta^T E_t\{\widehat{b}_T + \widehat{p}_T - \widehat{i}_T\} = 0 \tag{3.12}$$

where $\widehat{b}_t \equiv log(\frac{B_t}{P_t}) - log(\frac{B}{P})$, together with the constraint on the nominal public bond yield

$$\widehat{i}_t + \rho \geq 0 \ \ \forall t \tag{3.13}$$

where $\rho \equiv -log(Q) = -log(\beta)$.

3.3.4 Market clearing

Market clearing conditions are imposed for each variety i, namely $Y_t(i) = C_t(i)$, and they imply in aggregate (in log-deviations from steady-state):

$$\widehat{y}_t = \widehat{c}_t \tag{3.14}$$

where $\widehat{y}_t = log(Y_t) - log(Y)$ with $Y_t \equiv [\int_0^1 Y_t(i)^{1-\frac{1}{\varepsilon}} di]^{\frac{\varepsilon}{\varepsilon-1}}$. Aggregate price dynamics follow the same path as in the standard basic NK model (e.g. Chapter 3 in Galí (2015)), namely

$$\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda \widehat{mc}_t \tag{3.15}$$

with $\lambda \equiv \frac{(1-\theta)(1-\beta\theta)}{\theta}\frac{1-\alpha}{1-\alpha+\alpha\varepsilon}$, where $\widehat{mc}_t \equiv (\psi_t - p_t) - (-\mu) = \widehat{\omega}_t - (a_t - \alpha\widehat{l}_t)$ is the log-deviation from steady-state of the average real marginal cost.

The market clearing condition on the labor market writes

$$L_t = \int_0^1 L_t(i) di,$$

and implies up to a first order log-linear approximation

$$(1-\alpha)\widehat{l}_t = \widehat{y}_t - a_t \tag{3.16}$$

On the bond market, government supply equals household demand.

[7]In considering this particular steady-state, we follow Eggertsson and Woodford (2004) and Nakata (2017), by implicitly assuming that structural parameters of the model and the shocks are such that the zero-inflation steady-state is optimal. From the analyses in Budianto (2019) and Andrade et al. (2018), however, we can see that this is may not always be the case.

3.3.5 Equilibrium

Combining representative household's consumption/saving decision (3.5) with the goods market clearing condition (3.14), we can summarize aggregate demand by

$$\widehat{y}_t = E_t\{\widehat{y}_{t+1}\} - \frac{1}{\sigma}(\widehat{i}_t - E_t\{\pi_{t+1}\}) + \frac{1}{\sigma}(1-\rho_z)z_t \ \ \forall t \tag{3.17}$$

Furthermore, aggregate price dynamics (3.15) combined with the goods market clearing condition (3.14), the labor market clearing condition (3.16) and the labor supply (3.6), can be used to summarize the supply-side of the model economy as follows

$$\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\widehat{y}_t + \lambda\frac{\tau}{1-\tau}\widehat{\tau}_t \ \ \forall t, \tag{3.18}$$

Finally, using (3.6), (3.16) and (3.14), we can write the flow-budget constraint of the public sector (3.11) also in terms of aggregate variables

$$\widehat{b}_t = \beta^{-1}\left[\widehat{b}_{t-1} - \pi_t - \left((1-\beta) + \frac{\bar{t}}{\bar{b}}\right)\left(\left(\sigma + \frac{1+\varphi}{1-\alpha}\right)\widehat{y}_t + \frac{1}{1-\tau}\widehat{\tau}_t\right)\right] + \widehat{i}_t \ \ \forall t \tag{3.19}$$

The tax rate τ_t and the policy rate $\widehat{i}_t$ are both chosen as endogenous responses to macroeconomic developments under the optimal policy mix.

3.3.6 Calibration

Parameters are calibrated as in Nakata (2017) where a similar extension of the basic NK model (but with positive public expenditures) is used to study optimal monetary-fiscal policy in the presence of ZLB constraints, namely $\beta = \frac{1}{1+0.075}$, $\varphi = 1$, $\theta = 0.75$, $\alpha = 0$, $\varepsilon = 10$, $\sigma = 1/6$, $\bar{b} = 2.4(0.5)$[8].

3.3.7 Steady-state Laffer curve

Proportional labor income taxes generate an endogenous Laffer-curve steady-state relation between the tax rate τ and tax revenues $\tau\frac{W}{P}L$. Specifically, a higher rate increases tax revenues when the current rate is below a certain threshold, but it reduces it when it is above it (figure 3.4). This is because an increase in the tax

[8] I use this calibration instead of the textbook calibration in Galí (2015), because it implies a nicely-shaped Laffer curve with a peak attained for a labor income tax rate of 50%. By constrast, the textbook calibration implies a Laffer curve (included in the Appendix 3.7.2 on page 132) very skewed to the right with a peak attained for a significantly higher labor income tax rate of 90%.

rate has two opposing effects on tax revenues. On the one hand, it has a direct positive effect at a given tax base $\frac{W}{P}L$. On the other hand however, it has an indirect negativ effect because it shrinks the tax base. This is because it decreases the after-tax wage (equation (3.4)) inducing households to work less (L declines), and hence produce less (Y declines) in equilibrium. As a result, labor income $\frac{W}{P}L$, which is a direct function of output, declines[9]. Up to a certain threshold value of the tax level $\bar{\tau}$, the former effect is stronger and an increase in the tax rate increases revenues. Beyond this threshold however, the latter effect prevails, and tax revenues decline.

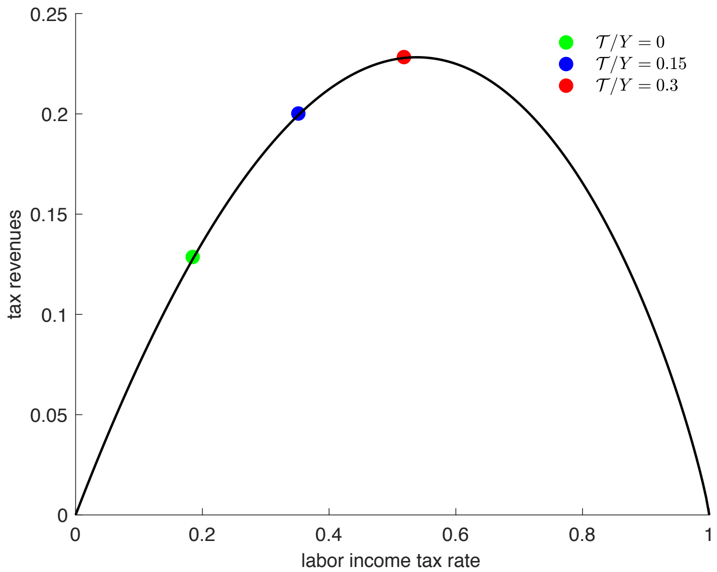


Figure 3.4: Steady-state Laffer curve

We call this steady-state revenue-maximizing tax rate *the long-run 'fiscal limit'*. As already pointed out, its key feature is that any further increase in the tax rate above this level $\bar{\tau}$ induces a decline in tax revenues in equilibrium. Otherwise stated, this tax rate is associated to the highest level of tax income that can be collected given the structural parameters of the economy. Using the zero-inflation steady-state relation implied by the government flow of funds (3.9)

$$\bar{\tau}\frac{W}{P}L = \frac{B}{P}\Big(1 - \beta\Big) + \mathcal{T}, \tag{3.20}$$

[9]The steady-state wage bill (hence, the tax base) equals $\frac{W}{P}L = (1-\alpha)\mathcal{M}^{-1}Y$.

we can see that this tax rate is also associated to the highest level of risk-free public debt and period real transfers that the government can service in equilibrium[10]. If the policymaker were to promise to repay higher such levels, it would not have the means to deliver on its promises. It would have to default, either outright or indirectly via inflation given that public debt is nominal[11].

As implied by the government flow of funds (3.20) above, high stable real long-run transfers $\mathcal{T}$ imply high real tax receipts $\tau\frac{W}{P}L$ needed to finance them. Thus, depending on $\mathcal{T}$, the economy may be far or close to its long-run fiscal limit. Figure 3.4 depicts the position of the economy on its long-run Laffer curve for three different values of real transfers: 0 (green), medium (blue), high (red), and mentions the associated steady-state transfers-to-output ratio in equilibrium. For a given real (inherited) debt level $\frac{B}{P}$, the higher the promised real long-run transfers, the higher the equilibrium steady-state transfers-to-output ratio, and the closer the economy to its 'fiscal limit'[12].

Our welfare analysis focuses on how the proximity of the economy to its long-run fiscal limit (i.e. to the peak of the Laffer curve in figure 3.4) affects the optimal stabilization policy over the business cycle[13]. The analysis concerns levels of real transfers that the policymaker can pay in equilibrium. The focus is thus on how promises *a priori* sustainable in the long-run may restrict the policy stabilization power at business cycle frequency by reducing the available fiscal space. We restrict attention to *long-run* tax rate levels on the increasing region of the curve.

3.4 Optimal policy design

The monetary-fiscal policy mix consists in the announcement of state-contingent plans for the nominal interest rate i_t and the tax rate τ_t. When taxes are distortionary, monetary and fiscal policies need to coordinate to engineer the optimal response to business cycle fluctuations[14]. In our welfare analysis, we study how

[10]Note that an upper bound on the left-hand-side of the equality, implies an upper bound on right-hand-side.

[11]Note that because debt is nominal, its real return can be eroded by inflation:

$$\bar{\tau}\frac{W}{P}L = \Pi^{-1}\frac{B}{P}\Big(1-\beta\Big) + \mathcal{T}$$

[12]Derivations used to generate figure 1 are included in the Appendix on page 131. Note that higher real levels of inherited debt also push economies to their fiscal limit.

[13]Importantly, if taxes were collected in a lump-sum fashion, the economy would not face any fiscal limit.

[14]This was not the case in the basic NK model with lump-sum taxes. As shown in Section 3.7.4 page 134, when taxes are collected in a lump-sum fashion, the optimal policy mix can be implemented (i) by having the monetary authority choose at each date the path of interest rates

the proximity to the fiscal limit affects the nature of this optimal coordination. In the first part of the analysis, we abstract from the existence of a ZLB on the policy rate, and in the second part we take it as well into consideration.

3.4.1 Optimal policy and the fiscal limit without a ZLB constraint

An increase in real transfers increases *a priori* the amount of distortionary taxes required to finance them, and hence widens long-run distortions[15]. As shown next, the policymaker can (partially) correct these inefficiences, and should take them into account while designing its optimal response to business cycle fluctuations. As the economy approaches its fiscal limit however, the welfare costs of doing so increase. To point this out, we split our analysis in two parts. We first study optimal policy when the economy is far from its fiscal limit, and we then look how results change in its proximity.

Fiscal pressures far away from the fiscal limit We first focus on small values of transfers (small 'fiscal pressures') in an environment where steady-state distorsions are small enough to conduct a welfare analysis following the standard textbook approach in Galí (2015), Section 5.3. To do so, we also consider a case with a smaller value of the long-run debt to output ratio (i.e. $\bar{b} = 0.1$) and a subsidizing scheme correcting for monopolistic market distorsions. As shown in the Appendix 3.7.5 on page 136, in this case the model economy is far away from its long-run fiscal limit. We assume the economy is in steady-state before the shock arises. The benevolent policymaker

$$\min_{\widehat{p}_t, \pi_t, \widehat{\widetilde{y}}_t, \widehat{i}_t, \widehat{\tau}_t, \widehat{b}_t} E_0 \sum_{t=0}^{\infty} \beta^t \Big[-\Lambda \widehat{\widetilde{y}}_t + \frac{1}{2}\Big(\pi_t^2 + \nu \widehat{\widetilde{y}}_t^2\Big)\Big] \text{ subject to}$$

$$\widehat{\widetilde{y}}_t = E_t\{\widehat{\widetilde{y}}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\}\Big) + \frac{1}{\sigma}(1-\rho_z) z_t, \quad \forall t \tag{3.21}$$

$$\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\Big(\sigma + \frac{\alpha+\varphi}{1-\alpha}\Big)\widehat{\widetilde{y}}_t + \lambda \frac{\tau}{1-\tau}\widehat{\tau}_t, \quad \forall t \tag{3.22}$$

needed to achieve optimal inflation and output (independently of fiscal developments), and (ii) by having the fiscal policy choose taxes and debt so as to satisfy its budgetary constraints given the (optimal) choices made by the monetary authority. This is because choices made by the monetary authority affect government's budget constraints by changing the paths of inflation, output (gap) and interest rates – details are given in the Appendix 3.7.3 on page 132.

[15] As shown by equation (3.50) on page 136 in the Appendix 33.7.5, the size of steady-state distorsions equals the tax rate. Hence, on the convergence path to the fiscal limit, the higher the taxes, the higher associated distorsions.

$$\widehat{b}_t = \beta^{-1}\left[\widehat{b}_{t-1} - \pi_t - \left((1-\beta) + \frac{\bar{t}}{\bar{b}}\right)\left(\left(\sigma + \frac{1+\varphi}{1-\alpha}\right)(\widehat{\widetilde{y}}_t + \widehat{y}^e_t) + \frac{1}{1-\tau}\widehat{\tau}_t\right)\right] +$$
$$+ \widehat{i}_t, \quad \forall t \tag{3.23}$$
$$\lim_{T\to\infty} \beta^T E_t\{\widehat{b}_T + \widehat{p}_T - \widehat{i}_T\} = 0, \quad \forall t \tag{3.24}$$
$$\left(\widehat{i}_t + \rho \geq 0\right), \quad \forall t \tag{3.25}$$
$$\widehat{p}_t = \pi_t + \widehat{p}_{t-1}, \quad \forall t \tag{3.26}$$

where $\Lambda \equiv \Phi\lambda/\varepsilon > 0$ with $\Phi = 1 - \frac{1-\tau}{\mathcal{M}}$ the steady-state distorsion (derived in the Appendix on page 136), $\nu \equiv \frac{\lambda}{\varepsilon}\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)$, $\widehat{r}^e_t = (1-\rho_z)z_t$, $\widehat{y}^e_t = 0$ and $\widehat{\widetilde{y}}_t \equiv \widehat{y}_t - \widehat{y}^e_t$ measuring the deviation of the welfare-relevant output gap from its steady-state level.

Optimality conditions write (alongside all equality constraints):

$$\begin{aligned}
\pi_t :& \pi_t - (\sigma\beta)^{-1}\lambda^1_{t-1} + \lambda^2_t - \lambda^2_{t-1} + \beta^{-1}\lambda^3_t - \lambda^6_t = 0\\
\widehat{\widetilde{y}}_t :& -\Lambda + \nu\widehat{\widetilde{y}}_t + \lambda^1_t - \beta^{-1}\widehat{\lambda}_{t-1} - \lambda\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\lambda^2_t +\\
& + \lambda^3_t\beta^{-1}\left((1-\beta) + \frac{\bar{t}}{\bar{b}}\right)\left(\sigma + \frac{1+\varphi}{1-\alpha}\right) = 0\\
\widehat{i}_t :& \frac{1}{\sigma}\lambda^1_t - \lambda^3_t + \lambda^5_t = 0, t \neq T, \quad \frac{1}{\sigma}\lambda^1_T - \lambda^3_T - \lambda^4_T + \lambda^5_T = 0\\
\widehat{\tau}_t :& -\lambda\frac{\tau}{1-\tau}\lambda^2_t + \beta^{-1}\left((1-\beta) + \frac{\bar{t}}{\bar{b}}\right)\frac{1}{1-\tau}\lambda^3_t = 0\\
\widehat{b}_t :& \lambda^3_t - \lambda^3_{t+1} = 0, \forall t \neq T, \quad \lambda^3_T + \lambda^4_T = 0\\
& \left(\widehat{i}_t + \rho\right)\lambda^5_t = 0, \quad \lambda^5_t \leq 0, \widehat{i}_t + \rho \geq 0\\
\widehat{p}_t :& \lambda^6_t - \beta\lambda^6_{t+1} = 0\\
& \lambda^4_T + \lambda^6_T = 0
\end{aligned}$$

Figure 3.6 shows that even in the absence of exogenous disturbances, a mix of monetary and fiscal policies is used to correct for long-run inefficiencies. Specifically, the first top panel shows that a positive deviation in the welfare relevant output gap is engineered under optimal policy, implying that a positive permanent correction is applied to the long-run level of output.

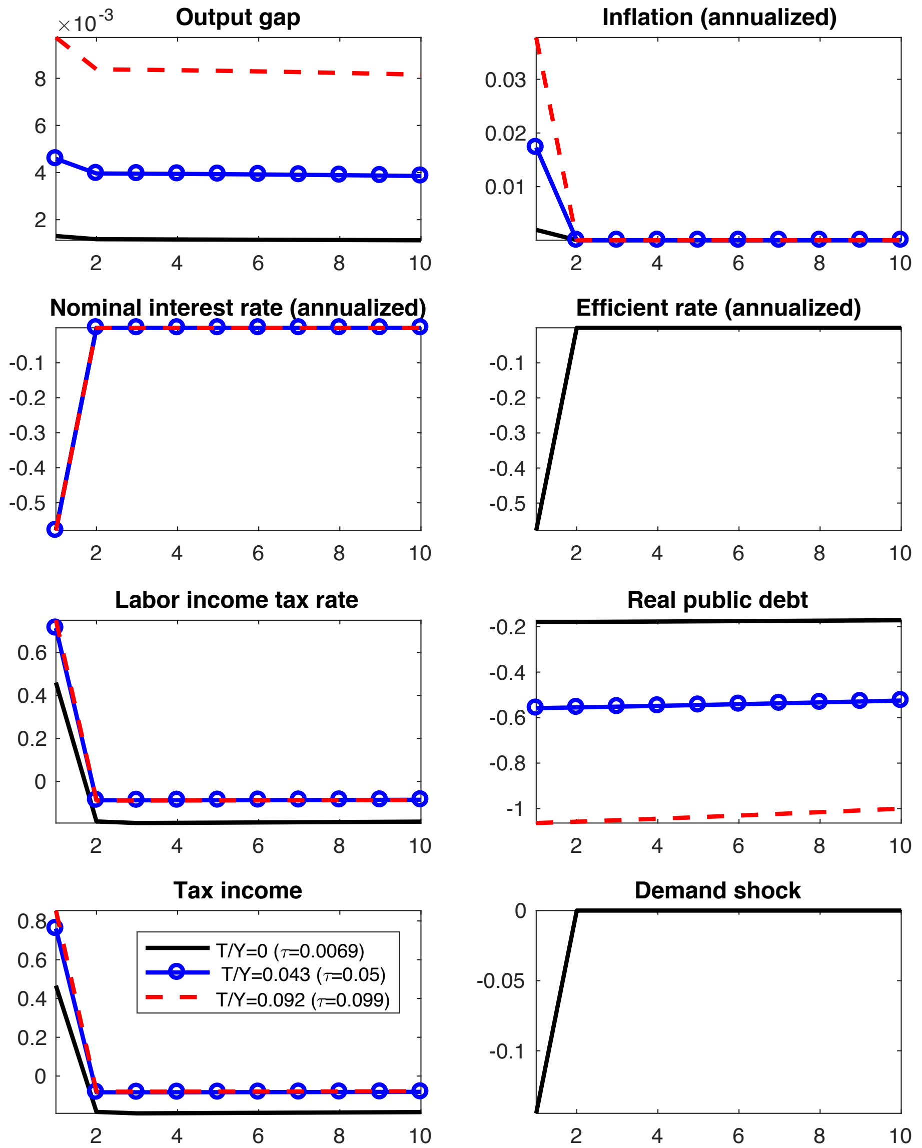


Figure 3.5: Dynamic responses to a transitory negative demand shock under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$ and no ZLB)
<u>Note:</u> Y-axis: $\%$ deviation from steady-state. X-axis: quarters

To understand how this is achieved, it is instructive to consider first the simple case with fully-flexible prices. When prices are fully flexible, monopolistic firms set prices as a constant markup over marginal costs (equal to $\mathcal{M}$), and hence the price setting equation (3.8) (in logs) is replaced by $0 = log\left(\frac{\mathcal{M}}{1-\alpha}\right) + \omega_t + l_t - y_t$. After combining it with the (log of the) production function (3.7) and the labor supply equation (3.6), we obtain that the equation describing the supply-side of the economy in the baseline model (3.22) is replaced in the case with flexible prices by

$$y_t = -\delta^{\tau y}\widehat{\tau}_t + \delta^y \tag{3.27}$$

where $\delta^{\tau y} \equiv \frac{1-\alpha}{\sigma(1-\alpha)+\alpha+\varphi}\frac{\tau}{1-\tau}$ and $\delta^y \equiv \frac{1-\alpha}{\sigma(1-\alpha)+\alpha+\varphi}\frac{\tau}{1-\tau}\Big(log(1-\tau) - log\left(\frac{\mathcal{M}}{1-\alpha}\right)\Big)$.

By setting the markup equal to 1 and the distorsionary tax rate equal to 0 in (3.27), we can derive the efficient level of output, namely the one which would arise with perfect competition and lump-sum taxes

$$y_t^e = \frac{(1-\alpha)log\big(1-\alpha\big)}{\sigma(1-\alpha)+\alpha+\varphi} \tag{3.28}$$

We can see from equations (3.27) and (3.28), that if the policymaker could permanently reduce the tax rate ($\widehat{\tau}_t < 0$) by the required amount, it could reestablish the equality between output and its efficient level, and eliminate any welfare losses. Figure 3.7 shows that this is possible if the policymaker simultaneously engineers an appropriate surprise increase in inflation so as to permanently erode the real value of debt, and hence its period real service cost[16]. This renders real public debt solvent despite the permanent decline in real tax revenues[17].

Things change when prices are sticky. In this case, if the policymaker wants to decrease the tax rate so as to push output above its initial steady-state level towards the efficient one, the surprise inflation required to restore the sustainability of public debt needs also to satisfy the price setting (supply) equation (3.22) (which was not the case when prices were fully flexible). Moreover, since variation in inflation is now costly in terms of welfare, such a policy may not even be optimal anymore. Indeed, figure 3.17 on page 138 in the Appendix shows that when price stickiness is low, tax rates decrease on impact, while the budget is balanced via

[16]Since public debt is nominal and non-state-contingent, the policymaker uses in fact unanticipated inflation as a lump-sum tax or transfer on financial wealth.

[17]As in Benigno and Woodford (2004) (page 288), when prices are fully-flexible, expected inflation and hence the evolution of nominal government debt are indeterminate. We thus follow this reference and we add to the assumed policy objective a small preference for inflation stabilization with no cost in terms of other objectives. In this case, the optimal policy will be then one that involved $E_t\{\pi_{t+1}\} = 0$ each period.

unexpected inflation as in the case with fully flexible prices. Things are however different for standard degrees of price stickiness[18].

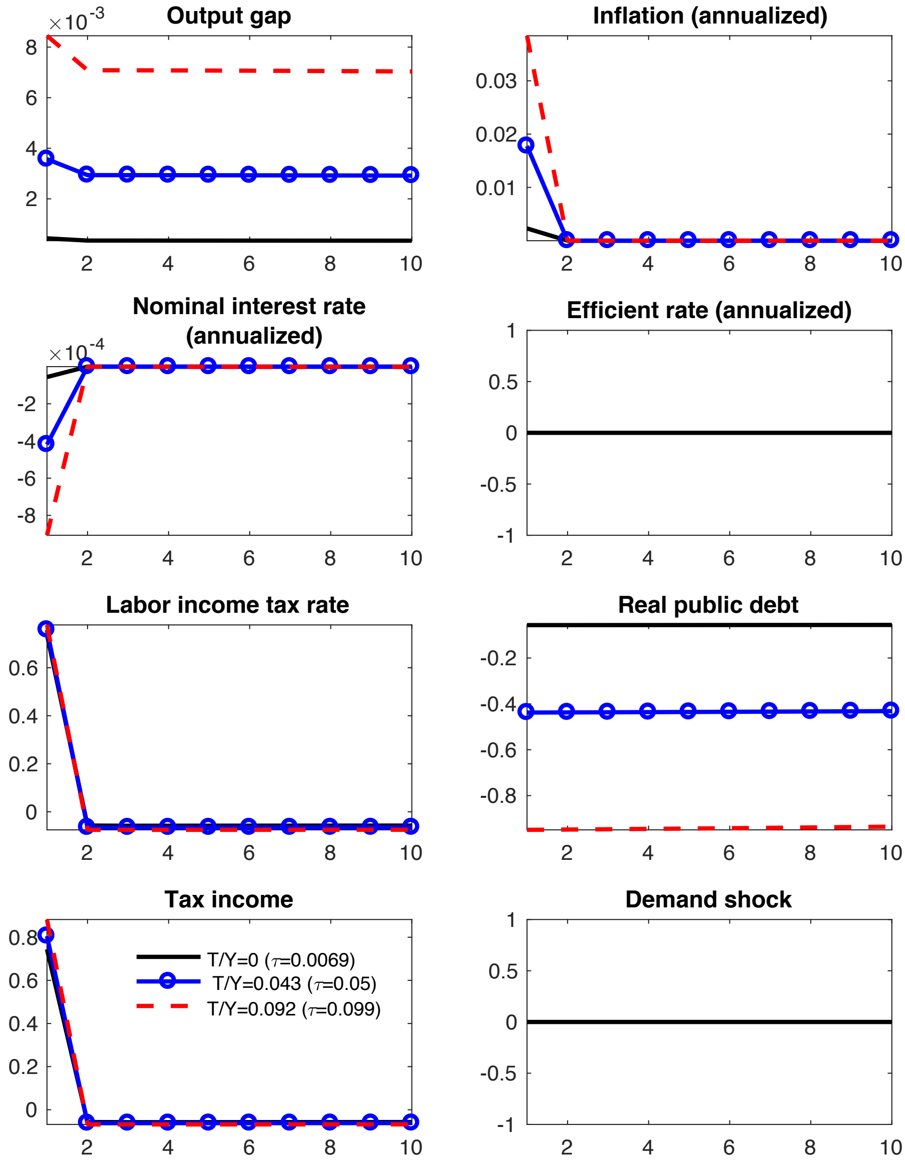


Figure 3.6: Optimal policy in the absence of shocks under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

[18]Conclusions reported here hold as well for the textbook calibration in Galí (2015).

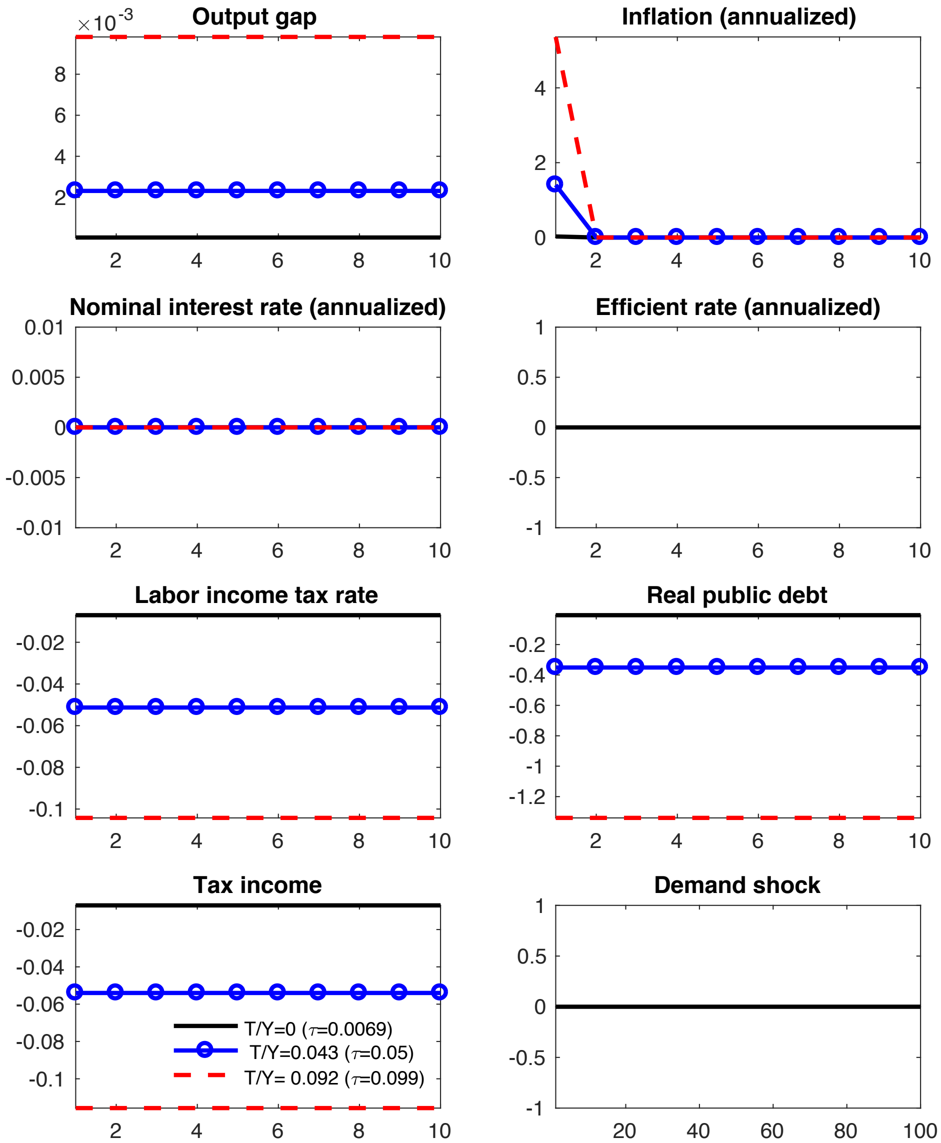


Figure 3.7: Optimal policy in the absence of shocks under optimal policy and flexible prices (small steady-state distorsions with $\bar{b} = 0.1$)

<u>Note:</u> Y-axis: % deviation from steady-state. X-axis: quarters

Returning to figure 3.6, we can see that, under optimal policy, long-run level of the tax rate is indeed lower than the initial one, whereas the long-run output is higher[19]. However, the initial decline in real debt which allows this is financed in a different way— specifically, it is engineered not only via unexpected inflation, but also via an initial increase in real taxes[20]. In terms of variation in policy instruments, this is achieved by a decline in the interest rate (second row, left panel) and an increase in the tax rate (third row, left panel)[21].

Everything else equal, the tax hike has a positive effect on the initial tax income (and hence a negative effect on real debt). However, it has also a negative effect on current output. With fully-flexible prices, such a policy would be suboptimal– recall that with fully-flexible prices, under optimal policy, output equals its efficient level at all times and hence there is no welfare trade-off between (lower) current and (higher) future outcome. When prices are sticky, there are two important differences. First, it is now costly to use (unexpected) inflation to finance the decline in real debt required to accommodate the permanent lower level of tax revenues (and higher output). Second, when aiming to use real taxes to engineer a decline in real debt, conditional on the degree of price stickiness being high enough, the policymaker does not necessary face a trade-off between (lower) current and (higher) future outcome. This is because it can now simultaneously use a decline in the interest rate to (temporary) increase demand and (more than) counteract the negative effects of the initial tax hike on output (this was not possible with flexible prices because the demand channel was not present in that case). Consequently, the policymaker can now temporary increase real taxes without having to incur a temporary decline in output. Indeed, for the standard degree of price stickiness considered in our analysis, the first top panel in figure 3.6 shows that the temporary positive effects on output are even stronger than the long-run ones under optimal policy[22]. Furthermore, since optimal policy supposes initially an increase in aggregate demand and a decline in aggregate supply, prices increase. This unexpected rise in inflation further contributes to the initial decline in real debt. Equilibrium inflation is however substantially lower than in the case

[19]Long-run levels refer to the (new) steady-state levels.

[20]The independent role played by the tax instrument in engineering a decline in the long-run level of debt can be observed by comparing results in figure 3.6 to the ones in the case with zero steady-state debt where inflation variations have no first order effect on debt (figure 3.24 on page 159 in the Appendix). In this case, as shown in the Appendix on page 153, variations in inflation and interest rates do not have any first order effects on real public debt developments. Thus, the increase in the labor income tax alone is the one triggering the decline in real public debt.

[21]Note that in the case of a negative demand shock, the increase in the tax rate is unrelated to the ZLB as in Eggertsson and Woodford (2004) or Nakata (2017).

[22]Note also that the magnitudes of the initial decline in the interest rate and of the tax rate hike are chosen so as to allow for a smooth path of the positive output gap.

of fully-flexible prices[23].

The results in the absence of exogenous disturbances help us understand the responses to the demand shock under optimal policy. Figure 3.5 depicts these dynamics for a negative shock, whereas figure 3.18 depicts them for a positive shock. By comparing these responses to the ones in figure 3.6 in the absence of shocks, we can see how the policy rate is used now both to respond to the exogenous demand disturbance and to correct for long-run inefficiencies.

Specifically, (it can be checked that) in the case of a negative demand shock the policy rate declines slightly more than one-for-one with the efficient rate in the first period, with the variation one-for-one with the efficient rate aimed to offset the effect of the demand shock, and the additional variation aimed to correct for long-run inefficiencies[24]. Consistently, in the case of a positive shock the policy rate increases slightly less in the first period. The relatively more accommodative stance than the one strictly warranted by the positive demand shock on impact (which would imply a one-for-one increase with the efficient rate) is again aimed at pushing output above its steady-state level in the first period (left top panel) and allowing for positive inflation in equilibrium (right top panel).

Furthermore, it can be observed in both cases how the tax rate is again used to additionally correct for long-run inefficiencies, but the magnitude of its variation now depends on the response of the policy rate to the shock. In particular, for a negative shock when the policy rate declines to offset its effect, the increase in the tax rate on impact is lower than in the absence of shocks. This is because the decline in the interest rate on impact has already negative effects on public debt, and hence the need for additional policy adjustments aimed at correcting long-run distorsions is lower. Consistently, in the case of a positive shock, the increase in the tax rate is larger than in the absence of the shock. This is explained by the need to accommodate the positive pressures on debt exerted by the rise in the policy rate. The relation between the policy rate and the tax rate can be better visually disentangled when comparing the effects of transitory versus persistent shocks.

[23]These results can be related to the ones in Schmitt-Grohe and Uribe (2004a) and Schmitt-Grohe and Uribe (2004b). The first paper finds that, when prices are flexible, a benevolent policy-maker which needs to finance an exogenous stream of government goods over the business cycle and has only distorsionary (labour income) taxes at its disposal, will use (unexpected) inflation to finance them, and will keep the labor income tax rate remarkably smooth. The second paper finds that a very small amount of price stickiness suffices to make the optimal inflation many times lower than that arising under full price flexibility. In this case the planner replaces unexpected inflation with standard debt and tax instruments.

[24]In the absence of steady-state distorsions, the policy rate optimally varies one-for-one with the efficient rate (Chapter 3 in Galí (2015)), and only the real rate reacts in equilibrium. Variations associated to steady-state distorsions are small, because the latter are small in the case analyzed in this section.

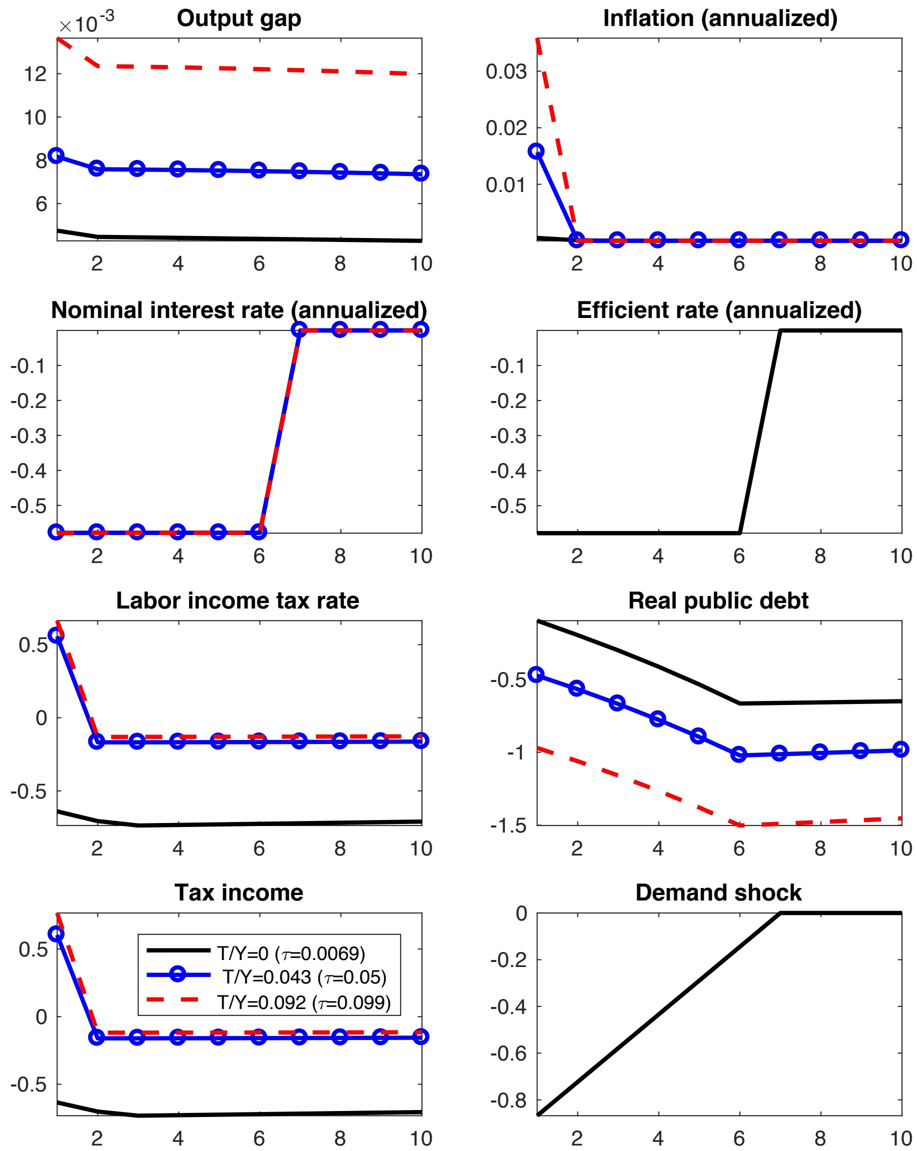


Figure 3.8: Dynamic responses to a six-period negative demand shock under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$ and no ZLB)
<u>Note:</u> Y-axis: $\%$ deviation from steady-state. X-axis: quarters

In particular, figure 3.8 shows how for a persistent negative demand shock where the policy rate (second row left panel) is kept below its steady-state value for longer (and, hence its effect on real debt is stronger) than in the case of a transitory shock (figure 3.5, second row left panel), the initial increase in the tax rate is lower (and even negative when steady-state distortions that need to be corrected are low). The converse is true for a positive demand shock. This can be observed by comparing the case of a six-period shock in figure 3.19 to the one of a transitory shock in figure 3.18.

Fiscal pressures in the proximity of the fiscal limit Now, we look at how real transfers affect optimal policy when their level is high, and in particular as they approach the economy to its fiscal limit. Since high transfer levels imply large steady-state distortions, we need to use the method developed by Benigno and Woodford (2004) to derive the optimal policy response. Again, the public sector sets its instruments i_t and τ_t so as to minimize representative household's welfare losses relatively to the efficient allocation subject to the constraints imposed by the functioning of the economy described by equations (3.17), (3.18), (3.13), (3.19) and (3.12). As shown in the Appendix 3.7.6 on page 139, the welfare losses equal up to a second order approximation

$$\mathcal{WL} = \frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left(Q_x^{22}\widetilde{y}_t^2 + Q_x^{11}\widehat{\tau}_t^2 + 2Q_x^{12}\widetilde{y}_t\widehat{\tau}_t + 2\xi_t^\tau \widehat{\tau}_t + q_\pi \pi_t^2 \right)$$

where $\widetilde{y}_t \equiv \widehat{y}_t - \widehat{y}_t^*$ is the welfare relevant output gap with $\widehat{y}_t^* \equiv -\frac{Q_\xi^{21}}{Q_x^{22}} z_t$ and $\xi_t^\tau \equiv \left(Q_\xi^{11} - \frac{Q_x^{12} Q_\xi^{21}}{Q_x^{22}} \right) z_t$ (a function of exogenous shocks). Coefficients Q are functions of structural parameters defined in the Appendix 3.7.6 on page 147. The constraints faced by the benevolent policymaker expressed in terms of $\widetilde{y}_t$ write:

$$(1)\ \ \widetilde{y}_t = E_t\{\widetilde{y}_{t+1}\} - \frac{1}{\sigma}(\widehat{i}_t - E_t\{\pi_{t+1}\}) + \frac{1}{\sigma}(1-\rho_z)z_t + E_t\{\Delta \widehat{y}_{t+1}^*\}\ \ \forall t,$$

$$(2)\ \ \pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\left(\widetilde{y}_t + \widehat{y}_t^*\right) + \lambda\frac{\tau}{1-\tau}\widehat{\tau}_t\ \ \forall t,$$

$$(3)\ \ \widehat{i}_t + \rho \geq 0\ \ \forall t,$$

$$(4)\ \ \widehat{b}_t = \beta^{-1}\left[\widehat{b}_{t-1} - \pi_t - \left((1-\beta) + \frac{\bar{t}}{\bar{\bar{b}}}\right)\left(\left(\sigma + \frac{1+\varphi}{1-\alpha}\right)(\widetilde{y}_t + \widehat{y}_t^*) + \frac{1}{1-\tau}\widehat{\tau}_t\right)\right] + \\ + \widehat{i}_t\ \ \ \forall t,$$

$$(5)\ \ \lim_{T\to\infty} \beta^T E_0\{\widehat{b}_T + \widehat{p}_T - \widehat{i}_T\} = 0,$$

$$(6)\ \ \widehat{p}_t = \pi_t + \widehat{p}_{t-1}\ \ \ \forall t.$$

The Lagrangian method gives the following optimality conditions describing the dynamics of the economy under optimal policy

$$
\begin{aligned}
\widetilde{y}_t :\ & Q_x^{22}\widetilde{y}_t + Q_x^{12}\widehat{\tau}_t + \lambda_t^1 - \beta^{-1}\lambda_{t-1}^1 - \lambda\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\lambda_t^2 \\
& + \beta^{-1}\left[(1-\beta) + \frac{\bar{t}}{\bar{\bar{b}}}\right]\left(\sigma + \frac{1+\varphi}{1-\alpha}\right)\lambda_t^4 = 0 \\
\pi_t :\ & q_\pi \pi_t - \frac{1}{\sigma\beta}\lambda_{t-1}^1 + \lambda_t^2 - \lambda_{t-1}^2 + \beta^{-1}\lambda_t^4 - \lambda_t^6 = 0 \\
\widehat{i}_t :\ & \frac{1}{\sigma}\lambda_t^1 + \lambda_t^3 - \lambda_t^4 = 0,\ \ t \leq T-1,\ \ \frac{1}{\sigma}\lambda_T^1 + \lambda_T^3 - \lambda_T^4 - \lambda_T^5 = 0, \\
\widehat{b}_t :\ & \lambda_t^4 - \lambda_{t+1}^4 = 0,\ \ t \leq T-1,\ \ \lambda_T^4 + \lambda_T^5 = 0, \\
\widehat{\tau}_t :\ & Q_x^{11}\widehat{\tau}_t + Q_x^{12}\widetilde{y}_t + \xi_t^\tau - \lambda\frac{\tau}{1-\tau}\lambda_t^2 + \frac{\beta^{-1}}{1-\tau}\left[(1-\beta) + \frac{\bar{t}}{\bar{\bar{b}}}\right]\lambda_t^4 = 0, \\
\widehat{p}_t :\ & \lambda_t^6 - \beta\lambda_{t+1}^6 = 0,\ \ t \leq T-1,\ \ \lambda_T^5 + \lambda_T^6 = 0, \\
& \lambda_t^3(\widehat{i}_t + \rho) = 0,\ \widehat{i}_t + \rho \geq 0,\ \lambda_t^3 \leq 0,
\end{aligned}
$$

together with the sequence of equality constraints above. λ_t^i are the Lagrangian multipliers associated to each of the six types of constraints at time t.

Figures 3.9 and 3.10 show that one implication of approaching the fiscal limit is that taxes cannot be used anymore so as to engineer an initial decline in public debt followed by a persistent stream of tax rates below average. Specifically, the higher the real transfers, and hence the closer the economy to its fiscal limit, the lower the initial optimal increase in tax income. This is in contrast to the case where the economy was very far from its fiscal limit, and higher real transfers were associated to higher optimal initial increases in taxes aimed at keeping public debt (and hence steady-state disorsions) persistently low in the following periods (figures 3.5 and 3.8). Monetary policy continues to be optimally used to fully-counteract the effect of the demand shock.

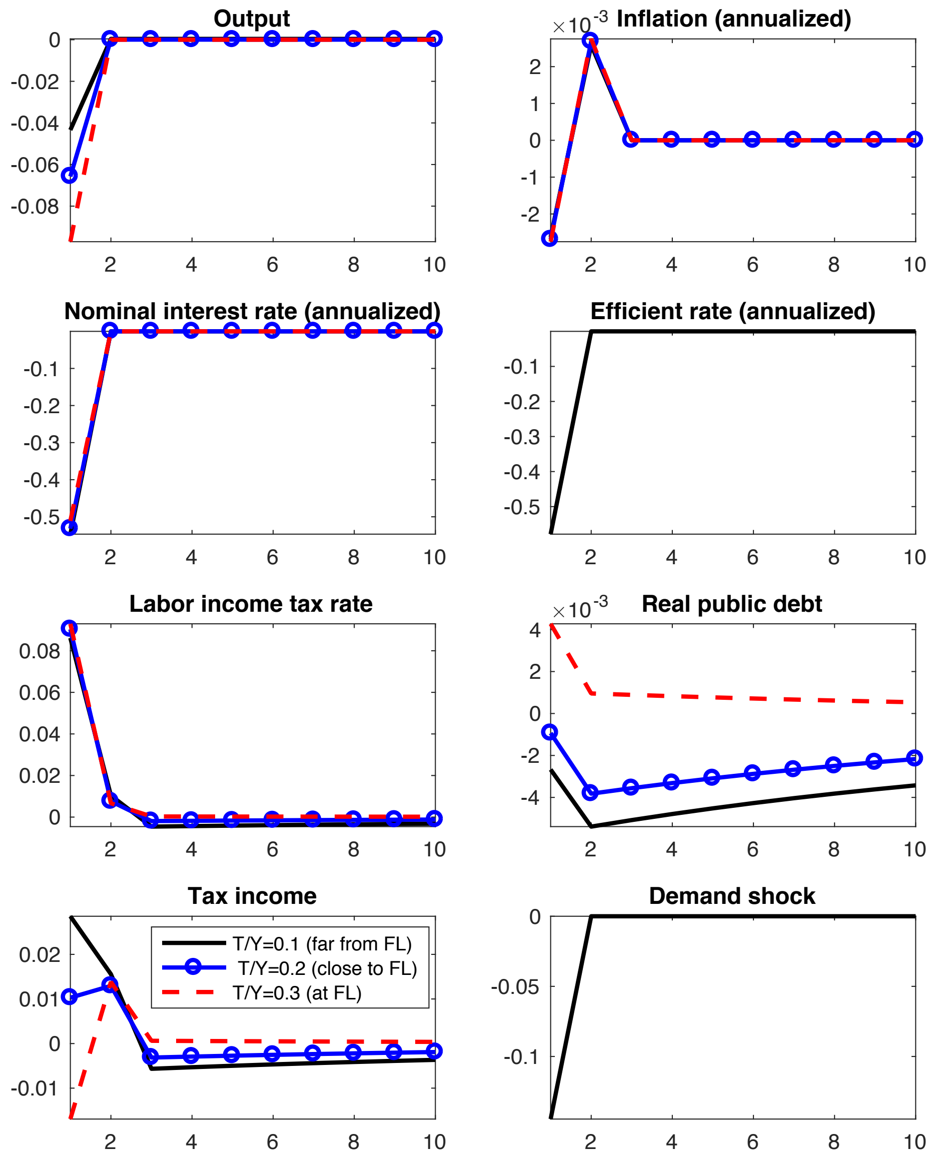


Figure 3.9: Dynamic responses to a transitory negative demand shock under optimal policy (large steady-state distorsions and no ZLB)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

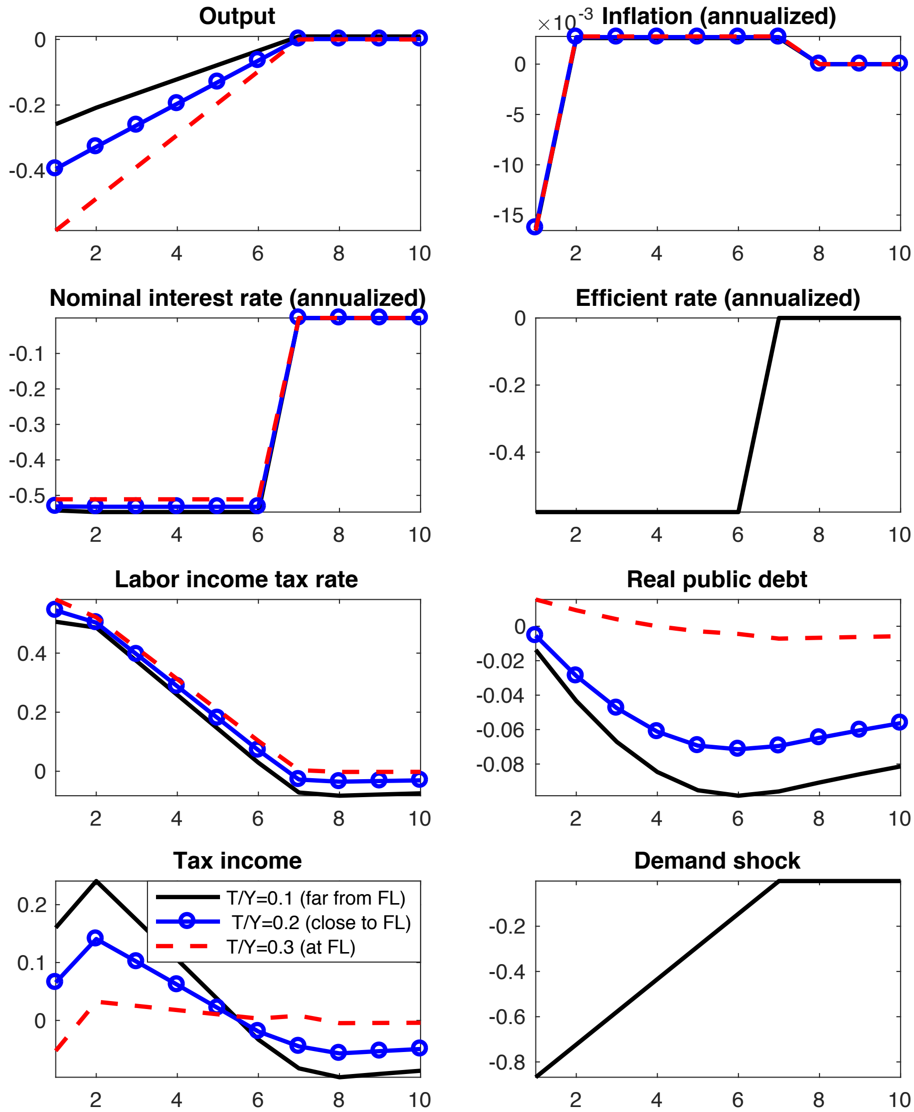


Figure 3.10: Dynamic responses to a six-period negative demand shock under optimal policy (case of large steady-state distorsions and no ZLB)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

3.4.2 Optimal policy and the fiscal limit with a ZLB constraint

As before, to disentangle the effect of the fiscal limit on optimal policy, we first study policy design when the economy is far away from it, and then, in a second step, in its proximity.

Optimal policy and the ZLB when the economy is far from fiscal limit We study this question in the simplest possible way, using the case with small steady-state distorsions described in section 3.4.1, with the difference that we now account as well for the ZLB. Figure 3.11 depicts the responses of the model economy under optimal policy in response to a transitory negative demand shock. The main insight is that when the ZLB constraints optimal policy, fiscal policy is used to compensate for the initial lack of ammunition of monetary policy.

Specifically, the tax rate is risen during the liquidity trap so as to counteract deflationary pressures, and to create fiscal space for a future decline in taxes. The additional fiscal space allows the policymaker to commit to a future tax cut (i.e. an expansionary fiscal policy) once the effects of the shock have dissipated. This 'promised' looser fiscal policy is aimed to engineer a future boom, and hence encourage present consumption. A similar accommodative policy stance is promised as well in terms of monetary policy, after the disturbance has ended, the policy rate remaining lower that would otherwise be chosen given the conditions prevailing at that time[25]. This 'history dependent' policy pattern has been already uncovered by Eggertsson and Woodford (2004)[26]. Figure 3.11 further shows how taxes retain their role in correcting long-run allocation inefficiencies (already pointed out when we derived optimal policy in the absence of the ZLB). Specifically, as the right panel on the third row in figure 3.11 shows, the policymaker designs their path such that real debt remains permanently lower in the aftermath of the shock– the larger transfers (and hence, distorsions), the larger the magnitude of the persistent decline in real debt. The permanent decline in real debt allows as before for a permanent decline in the long-run tax rate, and hence, for a permanent rise in long-run output.

Figure 3.12 shows results in the case of large transfer-to-output ratios as the economy approaches its fiscal limit. The latter constraints the level of accommodative policy that the policymaker can promise for the future (as hinted by the lower future inflation when the economy is at its fiscal limit). This explains the *a priori* counter-intuitive less inflationary policy mix chosen optimally as the

[25]Recall that in the absence of the ZLB (figure 3.5), the pattern of the tax rate was exclusively linked to its role in correcting long-run distorsions. Specifically, taxes were initially increased on impact, and then permanently kept bellow steady-state levels after the shock.

[26]We derive this result however for a labor income tax, whereas Eggertsson and Woodford (2004) derive it for a sales tax.

economy approaches its fiscal limit. Specifically, as the economy comes closer to its fiscal limit, the policymaker loses its ability to create fiscal space for a future boom by initially raising taxes.

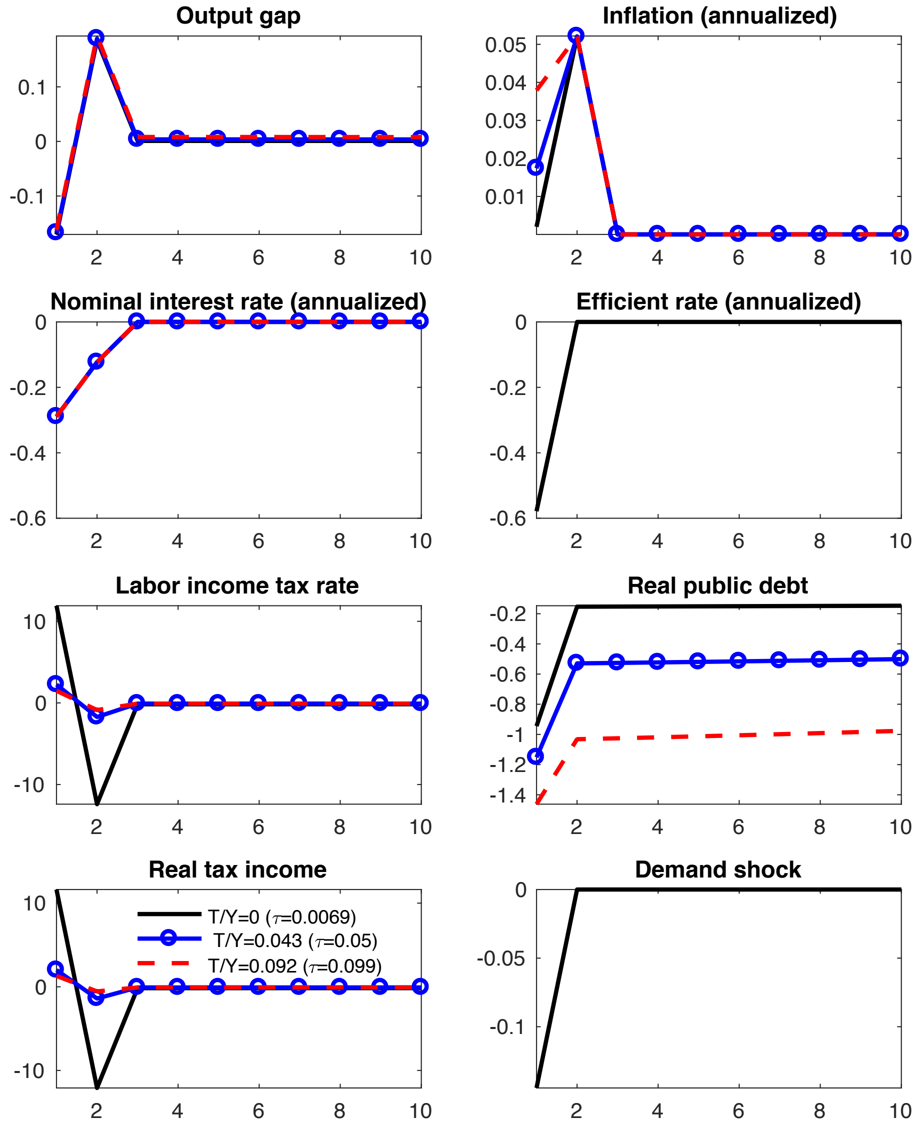


Figure 3.11: Dynamic responses to a transitory negative demand shock under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$ and ZLB)
<u>Note:</u> Y-axis: $\%$ deviation from steady-state. X-axis: quarters

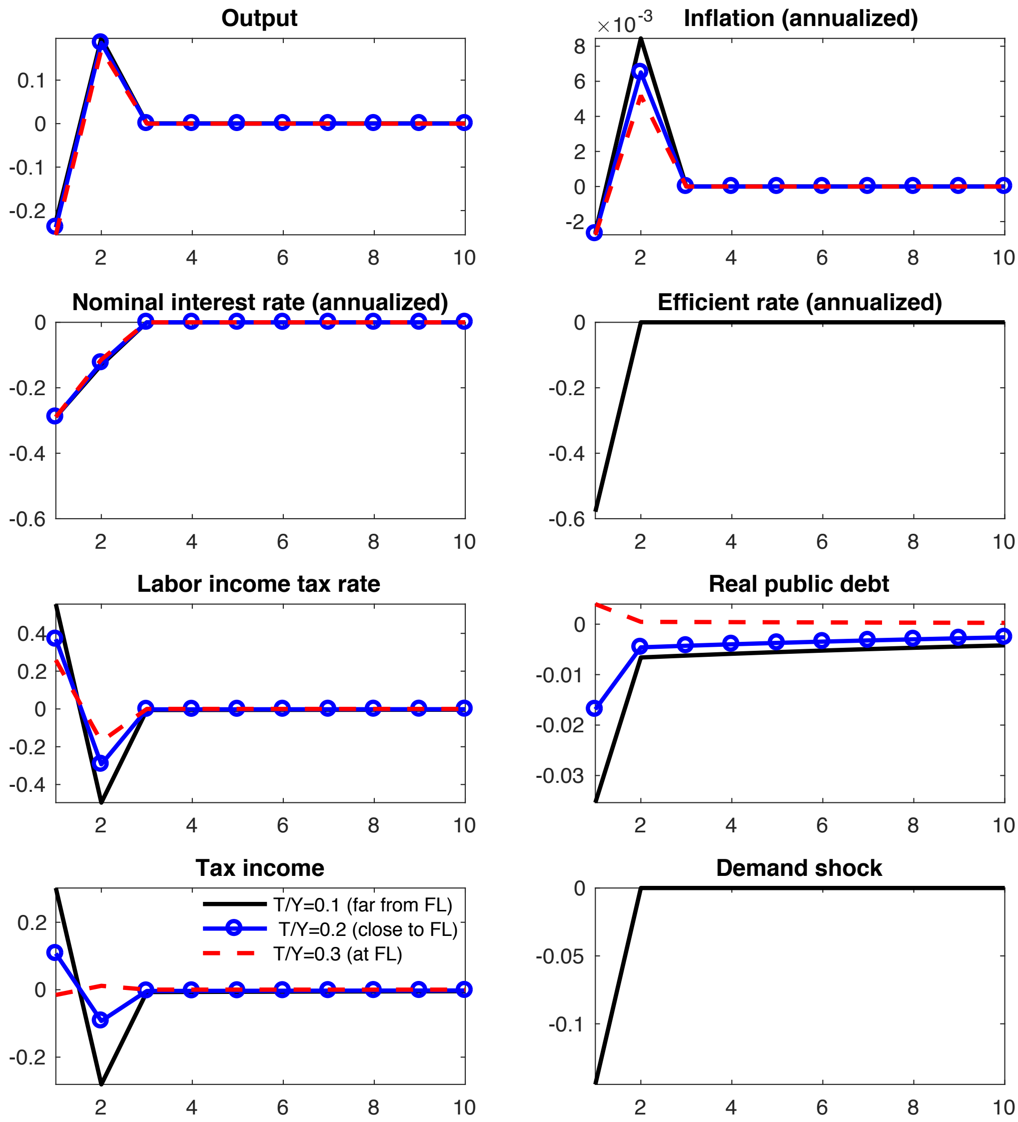


Figure 3.12: Dynamic responses to a transitory negative demand shock under optimal policy (large steady-state distorsions, convergence to fiscal-limit and ZLB)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

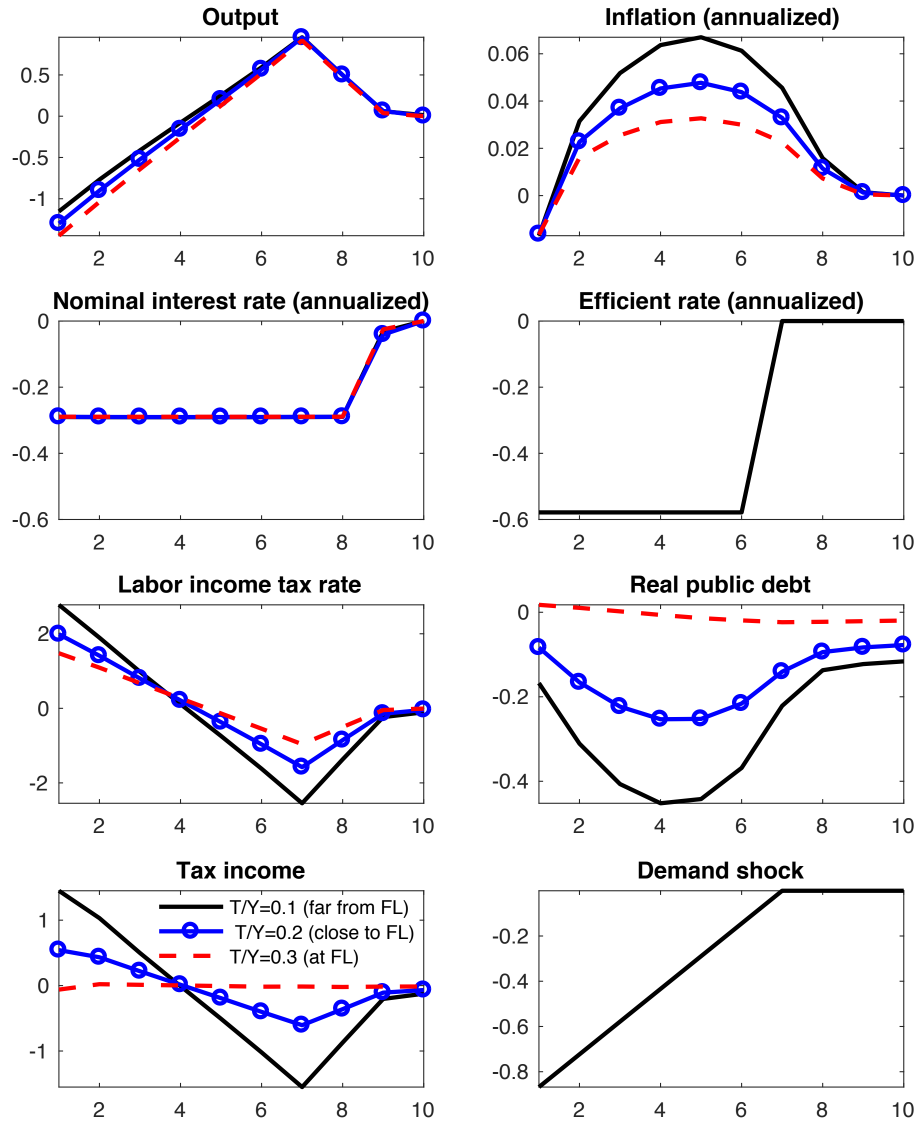


Figure 3.13: Dynamic responses to a six period negative demand shock under optimal policy (large steady-state distorsions, convergence to fiscal limit and ZLB)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

Thus, the higher real transfers (and hence, the closer the economy to its fiscal limit), the lower the initial increase in tax income and its future decline, and subsequently, the lower the promised future boom and future inflation under optimal policy. In the extreme case when the economy is at its fiscal limit, the initial optimal increase in the tax rate results in a decline in tax income, whereas its future decline results in an increase. Consistently, future inflation is the lowest in this extreme case. A shown in figure 3.11, such a pattern is not present when the economy is far away from the fiscal limit. In that case, higher real transfers are associated to higher inflation in equilibrium.

Therefore, apart from constraining the extent to which fiscal policy can be used to correct for steady-state distortions, the fiscal limit also constrains the extent to which tax policy can be used as a substitute to monetary policy when the latter is bounded by the ZLB. This result transpires even more clearly in the case of a persistent shock depicted in figure 3.13.

3.5 Conclusions

We studied in this paper how the zero lower bound and the fiscal limit affect the optimal monetary-fiscal response to demand shocks. We used for this purpose an extension of the basic NK model with an endogenous fiscal limit, and we abstracted away from both outright default on public debt and monetary financing. Our main result is that as the economy approaches its fiscal limit and monetary policy becomes constrained by the ZLB, dynamics under optimal policy become less inflationary in equilibrium. The main reason behind this finding is that the decrease in the fiscal space limits the extent of the future boom that the policymaker can promise so as to encourage current consumption (the "forward guidance").

Going forward, we want first to study how our results change when we allow for government expenditures as an additional policy instrument. Another more challenging future step is the analysis of the optimal monetary-fiscal policy mix beyond the fiscal limit. Using our current framework we can see that, excluding outright default on debt, one available policy option is an increase in the inflation target which would erode the real value of debt and of transfers (if transfers are made in nominal terms instead) up to the point where the peak of the Laffer Curve is reached again. But would this option be optimal? Namely, would this option which affects all agents in the economy be preferred to outright default on debt which affects only the elderly population (and is detrimental to the sovereign credit record)? And how does this choice depend on the presence of the ZLB which, by itself, may optimally require under certain conditions an increase in the inflation target? And if increasing the inflation target is the joint optimal solution to restore fiscal sustainability and deal with the ZLB, would optimal policy around a zero-inflation steady-state be the same as around a positive one? These are a few of the research questions that we aim to tackle in an extension of our setup in the future.

3.6 References

Andrade, Philippe, Jordi Galí, Herve Le Bihan, and Julien Matheron (2018): "The optimal inflation target and the natural rate of interest," National Bureau of Economic Research No. w24328.

Auerbach, Alan J., and Maurice Obstfeld (2005): "The case for open-market purchases in a liquidity trap," *American Economic Review*, 95(1), 110-137.

Bai, Yuting, and Eric M. Leeper (2017): "Fiscal stabilization vs. passivity," *Economics Letters*, 154 105-108.

Benigno, Pierpaolo, and Michael Woodford (2003): "Optimal monetary and fiscal policy: A linear-quadratic approach," NBER macroeconomics annual,18, 271-333.

Bi, Huixin (2012): "Sovereign default risk premia, fiscal limits, and fiscal policy," *European Economic Review*, 56(3), 389-410.

Bi, Huixin, and Eric M. Leeper (2013): "Analyzing fiscal sustainability," Bank of Canada Working Paper No. 2013-27.

Bianchi, Francesco, and Leonardo Melosi (2018): "The dire effects of the lack of monetary and fiscal coordination," *Journal of Monetary Economics*, 104, 1-22.

Bilbiie, Florin O. (2019): "Optimal forward guidance," *American Economic Journal: Macroeconomics*, 11.4, 310-45.

Bilbiie, Florin O., Tommaso Monacelli, and Roberto Perotti (2019): "Is government spending at the zero lower bound desirable?," *American Economic Journal: Macroeconomics*, 11.3, 147-73.

Budianto, Flora (2019): "Inflation Targets and the Zero Lower Bound," unpublished manuscript.

Burgert, Matthias, and Sebastian Schmidt (2014): "Dealing with a liquidity trap when government debt matters: Optimal time-consistent monetary and fiscal policy," *Journal of Economic Dynamics and control*, 47, 282-299.

Correia, Isabel, Emmanuel Farhi, Juan Pablo Nicolini, and Pedro Teles (2013): "Unconventional fiscal policy at the zero bound," *American Economic Review*, 103(4), 1172-1211.

Davig, Troy, Eric M. Leeper, and Todd B. Walker (2011): "Inflation and the fiscal limit," *European Economic Review*, 55(1), 31-47.

Davig, Troy, Eric M. Leeper, and Todd B. Walker (2010): "Unfunded liabilities? and uncertain fiscal financing," *Journal of Monetary Economics*, 57(5), 600-619.

Eggertsson, Gauti B., and Michael Woodford (2004): "Optimal monetary and fiscal policy in a liquidity trap," National Bureau of Economic Research No. w10840.

Eggertsson, Gauti (2001): "Real government spending in a liquidity trap," Princeton University, unpublished manuscript

Eggertsson, Gauti B., and Neil R. Mehrotra (2014): ”A model of secular stagnation,” NBER Working Paper No. 20547.

Eggertsson, Gauti B., Neil R. Mehrotra, and Jacob A. Robbins (2019): ”A model of secular stagnation: Theory and quantitative evaluation,” *American Economic Journal: Macroeconomics*, 11(1), 1-48.

Feldstein, Martin (2002): ”Commentary: Is there a role for discretionary fiscal policy?,” Rethinking Stabilization Policy, 151-162.

Ferrero, Giuseppe, Marco Gross, and Stefano Neri (2019): ”On secular stagnation and low interest rates: demography matters,” *International Finance*, 22(3), 262-278.

Galí, Jordi (2019): ”The Effects of a Money-Financed Fiscal Stimulus,” *Journal of Monetary Economics*.

Galí, Jordi (2020): ”Monetary policy and bubbles in a new Keynesian model with overlapping generations,” unpublished manuscript.

Galí, Jordi. (2013): ”Perceptions and Misperceptions of Fiscal Inflation: A Comment,” in A. Alesina and F. Giavazzi eds. Fiscal Policy after the Financial Crisis, The University Chicago Press, 2013, 299-305

Gnocchi, Stefano (2013): ”Monetary commitment and fiscal discretion: The optimal policy mix,” *American Economic Journal: Macroeconomics*, 5.2,187-216.

Gordon, Robert J (2017): *The rise and fall of American growth: The U.S. standard of living since the Civil War,* Princeton University Press.

Leeper, Eric M., and Campbell Leith (2016): ”Understanding Inflation as a joint monetary-fiscal phenomenon,” In Handbook of Macroeconomics, 2, 2305-2415.

Leeper, Eric M., and Todd B. Walker (2011): ”Fiscal limits in advanced economies,” *Economic Papers: A journal of applied economics and policy,* 30(1), 33-47.

Leeper, Eric M. (2013): ”Fiscal limits and monetary policy,” National Bureau of Economic Research, No. w18877.

Leeper, Eric M. (2016): ”Why central banks should care about fiscal rules,” *Sveriges Riksbank Economic Review*, 3:109-125.

Leeper, Eric M., and Tack Yun (2006): ”Monetary-fiscal policy interactions and the price level: Background and beyond,” International Tax and Public Finance, 13(4), 373-409.

Nakata, Taisuke (2017): ”Optimal government spending at the zero lower bound: A non-Ricardian analysis,” *Review of Economic Dynamics*, 23, 150-169.

Nakata, Taisuke (2016): ”Optimal fiscal and monetary policy with occasionally binding zero bound constraints,” *Journal of Economic Dynamics and control*, 73, 220-240.

Rouzet, Dorothee, Aida Caldera Sánchez, Theodore Renault, and Oliver Roehn (2019): ”Fiscal challenges and inclusive growth in ageing societies,” OECD, September 2019.

Schmidt, Sebastian (2013): ”Optimal monetary and fiscal policy with a zero bound on nominal interest rates,” *Journal of Money, Credit and Banking*, 45(7), 1335-1350.

Schmitt-Grohé, Stephanie, and Martin Uribe (1997): ”Balanced-budget rules, distortionary taxes, and aggregate instability,” *Journal of political economy*, 105(5), 976-1000.

Schmitt-Grohé, Stephanie, and Martin Uribe (2007): ”Optimal simple and implementable monetary and fiscal rules,” *Journal of monetary Economics*, 54(6), 1702-1725.

Schmitt-Grohé, Stephanie, and Martin Uribe (2004a): ”Optimal fiscal and monetary policy under imperfect competition,”*Journal of Macroeconomics*, 26(2), 183-209.

Schmitt-Grohé, Stephanie, and Martin Uribe (2004b): ”Optimal fiscal and monetary policy under sticky prices,” *Journal of economic Theory*, 114(2), 198-230.

Summers, Lawrence H. (2014): ”U.S. economic prospects: Secular stagnation, hysteresis, and the zero lower bound,” *Business Economics,* 49(2):65-73

Tulip, Peter (2014): ”Fiscal Policy and the Inflation Target,” *International Journal of Central Banking*

Werning, Ivan (2011): ”Managing a liquidity trap: Monetary and fiscal policy,” National Bureau of Economic Research, No. w17344.

Woodford, Michael (2003): *Interest and Prices: Foundations of a theory of monetary policy,* Princeton University Press.

3.7 Appendix

Relation between efficient rate and demand preference shocks

I assume the exogenous demand disturbance is of a similar nature as the one in Galí (2015), Section 5.4. Specifically, the efficient rate remains constant to its steady-state level $\rho > 0$ up to (and including) period 0. In period 1 it unexpectedly drops to $-\rho < 0$ (in the case of a negative shock), or it unexpectedly increases to $3\rho < 0$ (in the case of a positive shock), and remains at that level for one period or for six periods. Afterwards it takes again its steady-state value $\rho > 0$. Once the unexpected change in period 1 occurs, the subsequent path of the natural rate is assumed to be known with certainty by all agents.

The efficient rate equals in the model $r_t^e = \rho + z_t - E_t\{z_{t+1}\}$, and hence its deviation from steady-state equals $\widehat{r}_t^e = z_t - E_t\{z_{t+1}\}$.

- In the case of a one period shock $z_t = \widehat{r}_t^e$, where $\widehat{r}_t^e = -\rho - \rho = -2\rho$ for a negative shock, and $\widehat{r}_t^e = 3\rho - \rho = 2\rho$ for a positive shock.
- In the case of a six period shock, for the negative shock $\widehat{r}_6^e = -2\rho$ implies $z_6 = -2\rho$. Furthermore, $\widehat{r}_5^e = z_5 - z_6 = -2\rho$ implies $z_5 = \widehat{r}_5^e + z_6 = -4\rho$. Similarly, $z_4 = -6\rho$, $z_3 = -8\rho$, $z_2 = -10\rho$, $z_1 = -12\rho$.
- In the case of a six period shock, for the positive shock $\widehat{r}_6^e = 2\rho$ implies $z_6 = 2\rho$. Furthermore, $\widehat{r}_5^e = z_5 - z_6 = 2\rho$ implies $z_5 = \widehat{r}_5^e + z_6 = 4\rho$. Similarly, $z_4 = 6\rho$, $z_3 = 8\rho$, $z_2 = 10\rho$, $z_1 = 12\rho$.

3.7.1 Deterministic steady-state

Household's behaviour is described by

$$C^\sigma L^\varphi = (1-\tau)\frac{W}{P} \tag{3.29}$$

$$Q = \beta \tag{3.30}$$

$$C + Q\frac{B}{P} = Q + (1-\tau)\frac{W}{P}L + \frac{Div}{P} + \mathcal{T}, \tag{3.31}$$

the one of firms by

$$Y = AL^{1-\alpha} \tag{3.32}$$

$$1 = \frac{\mathcal{M}}{1-\alpha}\frac{W}{P}\frac{L}{Y} \tag{3.33}$$

and the flow budget constraint of the public sector by

$$Q\frac{B}{P} = \frac{B}{P} + \left(\mathcal{T} - \tau\frac{W}{P}L\right) \quad \text{or,} \tag{3.34}$$

$$Q\bar{b} = \bar{b} + \left(\bar{t} - \tau\frac{W}{P}\frac{L}{Y}\right) \tag{3.35}$$

Relations (3.33) and (3.35) imply that the tax rate needed to finance $\bar{b}$ and $\bar{t}$ equals:

$$\tau = \frac{\mathcal{M}\left[\bar{b}(1-\beta) + \bar{t}\right]}{1-\alpha} \tag{3.36}$$

Furthermore, (3.29), (3.32), (3.33) and the goods market clearing condition $Y = C$, imply that the steady-state level of labor

$$L = \Big[(1-\tau)(1-\alpha)\mathcal{M}^{-1}A^{1-\sigma}\Big]^{\frac{1}{\varphi+\sigma+\alpha(1-\sigma)}} \tag{3.37}$$

is a decreasing function of τ (given in (3.36)) and structural parameters.

The steady-state values of all other values can now be computed using L in (3.38). In particular, output equals

$$Y = A\Big[(1-\tau)(1-\alpha)\mathcal{M}^{-1}A^{1-\sigma}\Big]^{\frac{1-\alpha}{\varphi+\sigma+\alpha(1-\sigma)}} \tag{3.38}$$

(note that it is also a decreasing function of τ) and tax revenues

$$\tau\frac{W}{P}L = \tau(1-\alpha)\mathcal{M}^{-1}Y = \tau(1-\alpha)\mathcal{M}^{-1}A\Big[(1-\tau)(1-\alpha)\mathcal{M}^{-1}A^{1-\sigma}\Big]^{\frac{1-\alpha}{\varphi+\sigma+\alpha(1-\sigma)}} \tag{3.39}$$

Note that tax revenues are a concave function of τ.

For future reference, note in (3.33) that monopolistic market power distorsions can be corrected by subsidizing firm employment at the rate ε^{-1}, and financing these subsidies with lump-sum taxes. In the context of our analysis, this would imply lower (net) real steady-state transfers to the household.

3.7.2 Steady-state Laffer curve with textbook calibration

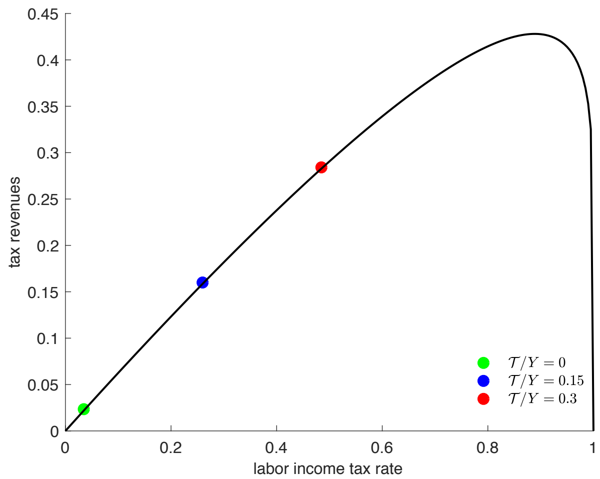


Figure 3.14: Steady-state Laffer curve with textbook calibration $\beta = 0.99$, $\varphi = 5$, $\theta = 0.75$, $\alpha = 0.25$, $\varepsilon = 9$, $\sigma = 1$, $\bar{b} = 2.4$ (Galí (2015))

3.7.3 Distortionary taxes and monetary-fiscal coordination

When characterizing the optimal policy mix with distortionary taxation, it is important to point out an important difference with respect to the benchmark case with lump-sum taxes. As shown in Section 3.7.4 page 134, when taxes are collected in a lump-sum fashion, the optimal policy mix can be implemented (i) by having the monetary authority choose at each date the path of interest rates needed to achieve optimal inflation and output (independently of fiscal developments), and (ii) by having the fiscal policy choose taxes and debt so as to satisfy its budgetary constraints given the (optimal) choices made by the monetary authority[27].

Subsequently, with lump-sum taxes, using the terminology of Leeper (1991), the optimal policy mix is always characterized by an *active* monetary stance (because to implement it, the monetary authority can set its instrument without paying attention to fiscal developments) and a *passive* fiscal stance (because, to implement it, the fiscal authority has to set its instruments to maintain the solvency of its budgetary constraints)[28]. Importantly, *this result is independent of whether the policy rate is constrained by the zero lower bound or not*. Thus, to implement *optimal policy* when taxes are collected in a lump-sum fashion, the monetary authority does not need to coordinate with the fiscal one[29]. This is in line with the optimality of an independent central-bank.

This conclusion changes however when taxes are proportional to labor income. Assume *optimal policy* could be implemented as before by having (i) the monetary authority choose the optimal dynamic paths of output, inflation and interest rates without paying attention to fiscal developments (i.e. via an 'active monetary' stance), and (ii) the fiscal authority choose the paths of tax rates and debt to accommodate this optimal choice. Using the New Keynesian Philips curve (constraint (2)), the welfare criterion can be written exclusively as a function of output and inflation. So, a priori, the monetary authority may choose an allocation that maximizes household's welfare without taking the tax rate into consideration. However, as put forward by the system of equations describing the non-fiscal block of the model (constraints (1), (2), (3)), for given paths of π_t, $\widehat{y}_t$ and i_t to be

[27]The choices made by the monetary authority affect government's budget constraints by changing the paths of inflation and interest rates.

[28]Leeper (1991) page 130: "I couch active and passive policy in terms of the constraints a policy authority faces. An active authority pays no attention to the state of government debt and is free to set its control variable as it sees fit. A passive authority responds to government debt (shocks). Its behaviour is constrained by private optimization and the active authority's actions". In our discussion the monetary authority sets its control variable (its policy rate instrument) in order to implement optimal policy, in Leeper (1991)'s analysis it was in order to pursue price stability by the means of a Taylor rule.

[29]By contrast, the fiscal authority coordinates with the monetary one when it adjusts its instruments to accommodate the choices of the latter.

implemented in equilibrium a specific unique path of tax rates τ_t is necessary[30]. And this specific path of the tax rate coupled with the chosen optimal output, implies a specific path for tax revenues. Subsequently, to passively accommodate the choices of the monetary authority, the fiscal authority is required to collect a certain amount of taxes at each date. However, this path of tax revenues, alongside the paths of interest rates and inflation is generally incompatible with the solvency of the government budget constraint. So, if the monetary policy would strictly choose its instrument to maximize the welfare criterion without paying attention to the implications of its choices for the solvency of government debt, it will generally make a choice that cannot be sustained in equilibrium[31]. This implies that such an allocation cannot be the optimal one.

It thus necessarily follows that, with proportional labor taxes, optimal policy cannot be implemented anymore by having the monetary authority take its decisions independently of the fiscal one (i.e. via an active monetary stance). Instead, it has to coordinate with the latter. In our welfare analysis, we study how the proximity to the fiscal limit affects the nature of this optimal coordination.

3.7.4 Optimal policy stance with lump-sum taxes

If taxes were collected in a lump-sum fashion, at any given level of period real transfers and long-run government debt, optimal policy would be the outcome of the following optimization problem

$$\min_{\{\pi_t,\widehat{y}_t,\widehat{i}_t,\widehat{t}_t^r,\widehat{b}_t,\widehat{p}_t\}} \frac{1}{2}\sum_{t=0}^{\infty}\beta^t\Big[-2\Phi\widetilde{y}_t+\frac{\epsilon}{\lambda}\pi_t^2+\Big(\sigma+\frac{\alpha+\varphi}{1-\alpha}\Big)\widetilde{y}_t^2\Big] \quad \text{subject to}$$

[30] Interestingly, if optimal policy could be implemented in this way with the monetary authority taking the lead and choosing the interest rate conditional on the fiscal authority optimally accommodating its choice (namely, via an 'active monetary/passive fiscal' stance), it could necessarily be also implemented via a 'passive monetary/active fiscal' stance where the fiscal authority took the lead and chose the tax rate conditional on the monetary policy following with the right choice of the interest rate. This was not the case with lump-sum taxes, since the fiscal authority could not use its instruments to affect the dynamics of output and inflation.

[31] This is reflected in the non-zero Lagrange multipliers associated with government's budgetary constraints

$$
\begin{aligned}
&(1)\ \ \widetilde{y}_t = E_t\{\widetilde{y}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\} - \widehat{r}^e_t\Big) \quad \forall t,\\
&(2)\ \ \pi_t = \beta E_t\{\pi_{t+1}\} + \kappa \widetilde{y}_t \quad \forall t,\\
&(3)\ \ \widehat{i}_t + \rho \geq 0 \quad \forall t,\\
&(4)\ \ \widehat{b}_t = \beta^{-1}\bigg[\widehat{b}_{t-1} - \pi_t - \Big((1-\beta) + \frac{\bar{t}}{\bar{b}}\Big)\widehat{t}^r_t\bigg] + \widehat{i}_t \quad \forall t,\\
&(5)\ \ \lim_{T\to\infty} \beta^T E_0\{\widehat{b}_T + \widehat{p}_T - \widehat{i}_T\} = 0,\\
&(6)\ \ \widehat{p}_t = \pi_t + \widehat{p}_{t-1} \quad \forall t.
\end{aligned}
$$

with $\Phi = 1 - \mathcal{M}^{-1}$ (due to steady-state market power distorsions), $\widehat{t}^r_t$ the log deviation from steady-state of long-run real lump-sum taxes, $\kappa \equiv$ and $\widehat{r}^e_t \equiv$. The Lagrangian method gives the following optimality conditions

$$\pi_t : \ \frac{\epsilon}{\lambda}\pi_t - \frac{1}{\sigma\beta}\lambda^1_{t-1} + \lambda^2_t - \lambda^2_{t-1} + \lambda^4_t\beta^{-1} - \lambda^6_t = 0 \tag{3.40}$$

$$\widetilde{y}_t : \ -\Phi + \Big(\sigma + \frac{\alpha+\varphi}{1-\alpha}\Big)\widetilde{y}_t + \lambda^1_t - \beta^{-1}\lambda^1_{t-1} - \kappa\lambda^2_t = 0 \tag{3.41}$$

$$\widehat{i}_t : \ \frac{1}{\sigma}\lambda^1_t + \lambda^3_t - \lambda^4_t = 0, \ \ t \leq T-1, \ \ \frac{1}{\sigma}\lambda^1_T + \lambda^3_T - \lambda^4_T - \lambda^5_T = 0, \tag{3.42}$$

$$\widehat{t}^r_t : \ \lambda^4_t\beta^{-1}\Big((1-\beta) + \frac{\bar{t}}{\bar{b}}\Big) = 0, \tag{3.43}$$

$$\widehat{b}_t : \ \lambda^4_t - \lambda^4_{t+1} = 0, \ \ t \leq T-1, \ \ \lambda^4_T + \lambda^5_T = 0, \tag{3.44}$$

$$\widehat{p}_t : \ \lambda^6_t - \beta\lambda^6_{t+1} = 0, \ \ t \leq T-1, \ \ \lambda^5_T + \lambda^6_T = 0, \tag{3.45}$$

$$\lambda^3_t(\widehat{i}_t + \rho) = 0, \ \widehat{i}_t + \rho \geq 0, \ \lambda^3_t \leq 0, \tag{3.46}$$

together with the constraints entering the optimization problem, where λ^i_t denote the Lagrangian multipliers associated to each of the six constraints.

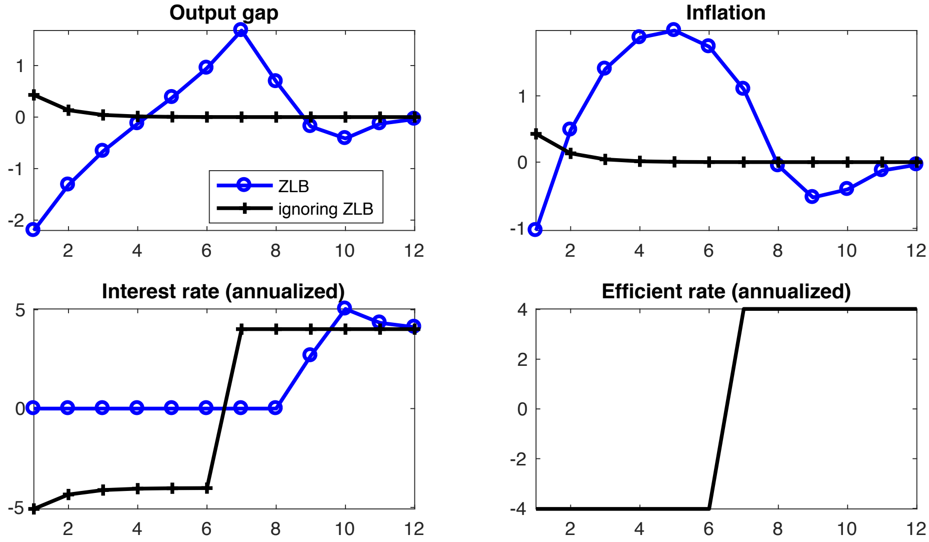


Figure 3.15: Dynamic responses to a six period negative demand shock under optimal policy (lump-sum taxes, market-power distorsions)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

Note that the Lagrange multipliers associated with the government flow of funds and its solvency λ_t^4, λ_t^5, λ_t^6 are zero at all dates. Consequently, the choices of taxes and debt do not constrain optimal allocation. Thus, in this case, the optimal monetary-fiscal policy mix is characterized by the monetary authority choosing at each date the interest rates needed to achieve optimal inflation and output gap, and by the fiscal policy choosing taxes and debt so as to satisfy its flow of funds and solvency constraints given the choices made by the monetary authority[32].

Subsequently, when taxes are collected in a lump-sum fashion the optimal policy mix regime is characterized by an *active* monetary stance and a *passive* fiscal stance in the sense of Leeper (1991). Importantly, *this result is independent on whether the policy rate is constrained by the zero lower bound or not*.

Importantly also, in a world with lump-sum taxes there is no fiscal limit. Thus, the debt burden and pre-committed real transfers do not constrain the stabilization power of policymakers. Subsequently, in the context of aging populations, only the ZLB may constrain it as a result of the decline in efficient real interest rates. Figure 3.15 makes this point by comparing the dynamic responses under optimal policy in the presence and absence of the ZLB on the policy rate.

[32]Note that the paths of taxes and debt associated to the optimal allocation are not unique, so the fiscal policy has just to pick one of these possible combinations *so as to ensure the sustainability of public debt*

3.7.5 Steady-state distorsion

The steady-state distorsion Φ is implicitly defined by

$$-\frac{U_l}{U_c} = (1-\Phi)\frac{\partial Y}{\partial L} \Rightarrow L^{\varphi}C^{\sigma} = (1-\Phi)(1-\alpha)\frac{Y}{L} \tag{3.47}$$

Using the good market clearing condition $C = Y$ and the production function (3.32), we can express the distorsion in terms of steady-state labor as

$$\Phi = 1 - \frac{A^{\sigma-1}L^{\varphi+\sigma+\alpha(1-\sigma)}}{1-\alpha}, \tag{3.48}$$

and further, using the expression of steady-state labor in (3.38), exclusively in terms of the two (steady-state) distorsionary sources- market power and distorsionary taxes:

$$\Phi = 1 - \frac{1-\tau}{\mathcal{M}} \tag{3.49}$$

The higher the market power in the goods market, or the higher taxes, the larger the distorsion of the steady-state allocation. When market-power distorsions are corrected (by subsidizing employment at rate ε^{-1} and financing these subsidies with lump-sum taxes)

$$\Phi = \tau \tag{3.50}$$

In section 3.4.1 we use a version of the model where labor income taxes are the only source of steady-state distorsion. In that case we can conduct the welfare analysis under the assumption of a small steady-state distorsions for steady-state tax rates τ strictly lower than 10^{-2}. The region on the Laffer curve associated to such taxes is shown in figure 3.16 below between the green and red points.

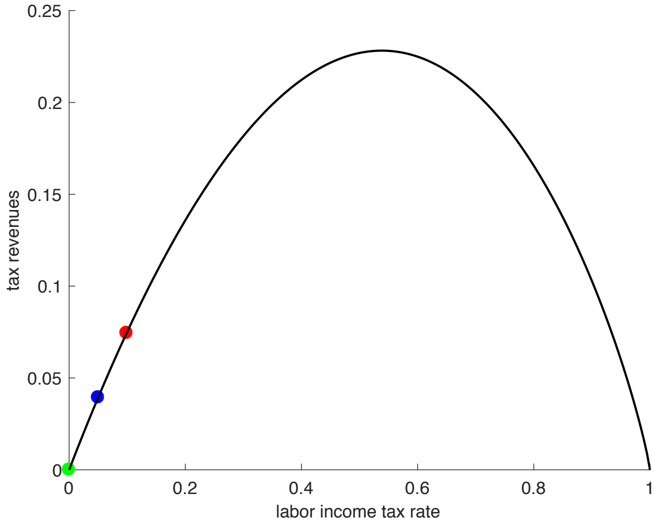


Figure 3.16: Steady-state Laffer curve and small distorsions

For the case of a strictly positive steady-state debt to output ratio I use $\bar{b} = 0.1$. In this case, $\frac{\mathcal{T}}{Y} = 0$ implies $\tau = 0.0069$, $\tau = 0.05$ implies $\frac{\mathcal{T}}{Y} = 0.043$, $\tau = 0.099$ implies $\frac{\mathcal{T}}{Y} = 0.092$. I also consider the case of a zero steady-state level of debt. Since I need in this version the same level of steady-state distorsions as in the one with positive debt, and the steady-state distorsions depend on the steady-state distorsionary tax rate, I consider the same steady-state tax rate levels as before with their associated transfer-to-output ratios, namely $\tau = 0.0069$ with $\frac{\mathcal{T}}{Y} = 0.0069$, $\tau = 0.05$ with $\frac{\mathcal{T}}{Y} = 0.05$ and $\tau = 0.099$ with $\frac{\mathcal{T}}{Y} = 0.099$[33].

[33]To compute these values I used $\frac{\mathcal{T}}{Y} = (1 - \alpha)\tau$.

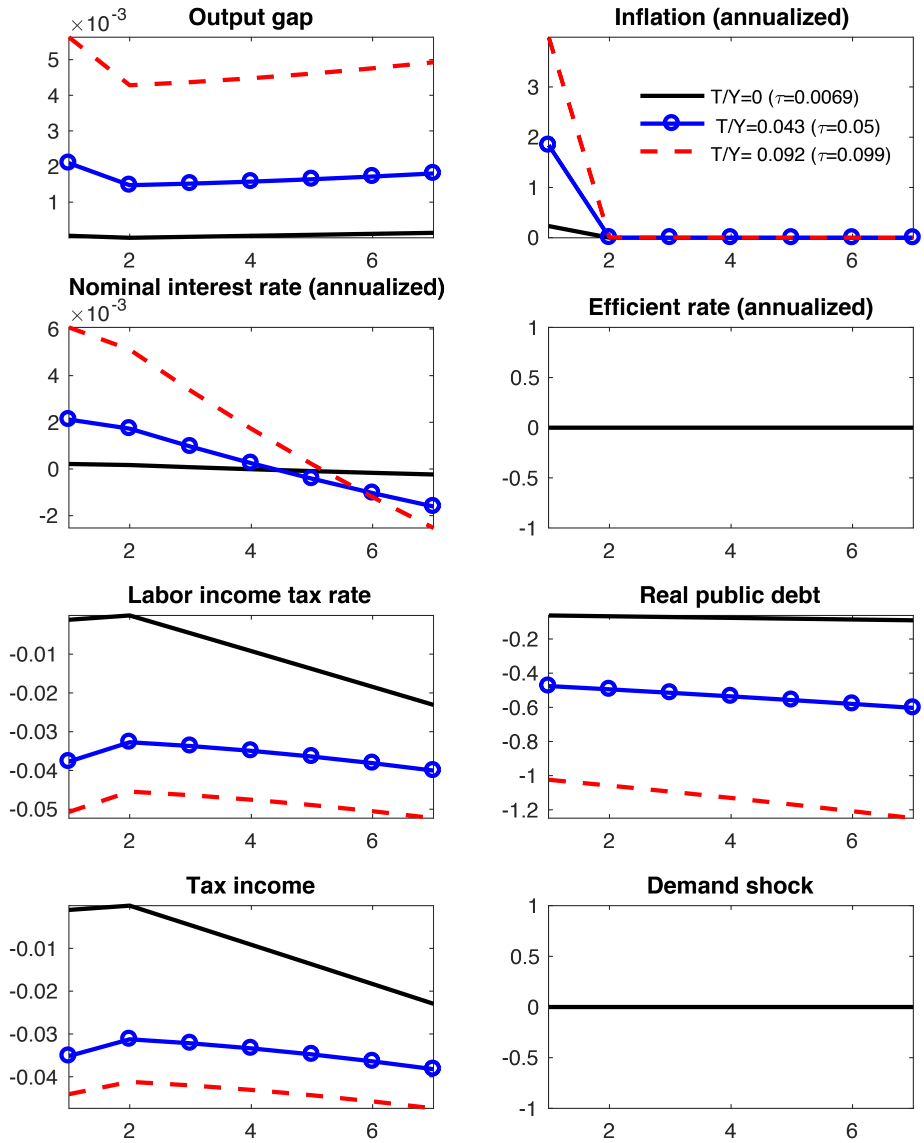


Figure 3.17: Optimal policy in the absence of shocks and low price stickiness (small steady-state distorsions with $\bar{b} = 0.1$, $\theta = 0.001$)

<u>Note:</u> Y-axis: % deviation from steady-state. X-axis: quarters

3.7.6 Welfare criterion

A second order approximation around steady-state to the period-utility of the representative household,

$$U(C_t, L_t, Z_t) = \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi}\right) Z_t, \tag{3.51}$$

yields:

$$\begin{aligned} U_t - U \approx\; & U_c C \frac{\Delta C_t}{C} + U_l L \frac{\Delta L_t}{L} + \frac{1}{2} U_{cc} C^2 \left(\frac{\Delta C_t}{C}\right)^2 + \frac{1}{2} U_{ll} L^2 \left(\frac{\Delta L_t}{L}\right)^2 + \\ & + U_{cz} C \frac{\Delta C_t}{C} \frac{\Delta Z_t}{Z} + U_{lz} L \frac{\Delta L_t}{L} \frac{\Delta Z_t}{Z} + t.i.p. \end{aligned}$$

where $t.i.p.$ are terms independent of policy. Using further the goods market clearing condition $C_t = Y_t$, we get:

$$\begin{aligned} \frac{U_t - U}{U_c C} \approx\; & \frac{\Delta Y_t}{Y} + \frac{U_l L}{U_c C} \frac{\Delta L_t}{L} + \frac{1}{2} \frac{U_{cc} C^2}{U_c C} \left(\frac{\Delta Y_t}{Y}\right)^2 + \frac{1}{2} \frac{U_{ll} L^2}{U_c C} \left(\frac{\Delta L_t}{L}\right)^2 + \\ & + \frac{\Delta Y_t}{Y} \frac{\Delta Z_t}{Z} + \frac{U_l L}{U_c C} \frac{\Delta L_t}{L} \frac{\Delta Z_t}{Z} + t.i.p. \\ \approx\; & \frac{\Delta Y_t}{Y}\left(1 + \frac{\Delta Z_t}{Z}\right) + \frac{U_l L}{U_c C} \frac{\Delta L_t}{L}\left(1 + \frac{\Delta Z_t}{Z}\right) + \frac{1}{2} \frac{U_{cc} C^2}{U_c C} \left(\frac{\Delta Y_t}{Y}\right)^2 + \\ & + \frac{1}{2} \frac{U_{ll} L^2}{U_l L} \frac{U_l L}{U_c C} \left(\frac{\Delta L_t}{L}\right)^2 + t.i.p. \end{aligned}$$

Given the utility specification in (3.51), $\frac{U_{cc} C^2}{U_c C} = -\sigma$ and $\frac{U_{ll} L^2}{U_l L} = \varphi$, and hence:

$$\begin{aligned} \frac{U_t - U}{U_c C} \approx\; & \frac{\Delta Y_t}{Y}\left(1 + \frac{\Delta Z_t}{Z}\right) + \frac{U_l L}{U_c C} \frac{\Delta L_t}{L}\left(1 + \frac{\Delta Z_t}{Z}\right) - \frac{\sigma}{2} \left(\frac{\Delta Y_t}{Y}\right)^2 + \\ & + \frac{\varphi}{2} \frac{U_l L}{U_c C} \left(\frac{\Delta L_t}{L}\right)^2 + t.i.p. \end{aligned}$$

Let Φ denote the size of the steady-state distorsion defined by $MRS = (1 - \Phi) MPL$, with $MRS \equiv -\frac{U_l}{U_c}$ the steady-state marginal rate of substitution and $MPL \equiv (1-\alpha) Y/L$ the steady-state marginal product of labor[34]. The expression

[34] $\Phi = 1 - \frac{1-\tau}{\mathcal{M}}$

above can be written in terms of this distorsion as:

$$\frac{U_t - U}{U_c C} \approx \frac{\Delta Y_t}{Y}\left(1 + \frac{\Delta Z_t}{Z}\right) - (1-\Phi)(1-\alpha)\frac{\Delta L_t}{L}\left(1 + \frac{\Delta Z_t}{Z}\right) - \frac{\sigma}{2}\left(\frac{\Delta Y_t}{Y}\right)^2 \\ - \frac{\varphi(1-\Phi)(1-\alpha)}{2}\left(\frac{\Delta L_t}{L}\right)^2 + t.i.p.$$

Using the second order approximation of relative deviations in terms of log deviations $\frac{X_t - X}{X} \approx \widehat{x}_t + \frac{1}{2}\widehat{x}_t^2$, we can approximate the expression above by:

$$\begin{aligned} \frac{U_t - U}{U_c C} &\approx \left(\widehat{y}_t(1+\widehat{z}_t) + \frac{1}{2}\widehat{y}_t^2\right) - (1-\Phi)(1-\alpha)\left(\widehat{l}_t(1+\widehat{z}_t) + \frac{1}{2}\widehat{l}_t^2\right) - \frac{\sigma}{2}\widehat{y}_t^2 \\ &\quad - \frac{\varphi}{2}(1-\Phi)(1-\alpha)\widehat{l}_t^2 + ||\mathcal{O}_t^3|| + t.i.p. \\ &\approx \widehat{y}_t(1+\widehat{z}_t) + \frac{1-\sigma}{2}\widehat{y}_t^2 - (1-\Phi)(1-\alpha)\widehat{l}_t(1+\widehat{z}_t) - \\ &\quad - \frac{1+\varphi}{2}(1-\Phi)(1-\alpha)\widehat{l}_t^2 + ||\mathcal{O}_t^3|| + t.i.p. \end{aligned}$$

A second order approximation of the labor market clearing condition (3.16) yields (e.g. Galí (2015), Chapter 4): $\widehat{l}_t = \frac{\widehat{y}_t - a_t + d_t}{1-\alpha}$ with $d_t \equiv \frac{\varepsilon}{2\Theta} var_i\{p_t(i)\}$, where $\Theta \equiv \frac{1-\alpha}{1-\alpha+\alpha\varepsilon}$, hence:

$$\frac{U_t - U}{U_c C} \approx \Phi\widehat{y}_t + \frac{1}{2}\left((1-\sigma) - \frac{(1-\Phi)(1+\varphi)}{1-\alpha}\right)\widehat{y}_t^2 - (1-\Phi)d_t + \\ + \frac{(1-\Phi)(1+\varphi)}{(1-\alpha)}a_t\widehat{y}_t + \Phi\widehat{y}_t\widehat{z}_t + t.i.p. + ||\mathcal{O}_t^3||$$

$$\frac{U_t - U}{U_c C} \approx \Phi\widehat{y}_t - \frac{1}{2}u_{yy}\widehat{y}_t^2 - \frac{1}{2}u_p var_i\{p_t(i)\} + \xi_t u_\xi \widehat{y}_t + t.i.p. + ||\mathcal{O}_t^3|| \quad (3.52)$$

with:

$$\begin{aligned} u_{yy} &\equiv \frac{(1-\Phi)(1+\varphi)}{1-\alpha} - (1-\sigma) \\ u_p &\equiv (1-\Phi)\frac{\varepsilon}{\Theta} \\ \xi_t u_\xi &\equiv \Phi\widehat{z}_t + \frac{(1-\Phi)(1+\varphi)}{(1-\alpha)}a_t \end{aligned}$$

Accordingly, a second order approximation to households' welfare losses expressed as a fraction of steady-state consumption equals:

$$\begin{aligned}\mathcal{WL} &= -E_0\sum_{t=0}^{\infty}\beta^t\biggl(\frac{U_t-U}{Uc\,C}\biggr)\\ &= E_0\sum_{t=0}^{\infty}\beta^t\biggl(-\Phi\widehat{y}_t+\frac{1}{2}u_{yy}\widehat{y}_t^2+\frac{1}{2}u_p var_i\{p_t(i)\}-\xi_t u_\xi\widehat{y}_t\biggr)+t.i.p.+||\mathcal{O}_t^3||\end{aligned}$$

which using $\sum_{t=0}^{\infty}\beta^t var_i\{p_t(i)\}=\frac{\theta}{(1-\beta\theta)(1-\theta)}\sum_{t=0}^{\infty}\beta^t\pi_t^2$ (Woodford (2003), Chapter 6) and the definition $u_\pi\equiv u_p\frac{\theta}{(1-\beta\theta)(1-\theta)}=(1-\Phi)\frac{\epsilon}{\lambda}$, can be written as:

$$\mathcal{WL}=E_0\sum_{t=0}^{\infty}\beta^t\biggl(-\Phi\widehat{y}_t+\frac{1}{2}u_{yy}\widehat{y}_t^2+\frac{1}{2}u_\pi\pi_t^2-\xi_t u_\xi\widehat{y}_t\biggr)+t.i.p.+||\mathcal{O}_t^3|| \tag{3.53}$$

Next, we follow the approach in Benigno and Woodford (2003) to eliminate the linear term $\Phi\widehat{y}_t$. We use for this purpose a second order approximation of the aggregate supply relation. The aggregate-supply relation can be written exactly as:

$$-log\biggl(\frac{1-\theta\Pi_t^{e-1}}{1-\theta}\biggr)=\frac{e-1}{1+e\omega}\Bigl(logK_t-logF_t\Bigr)\quad\text{where:} \tag{3.54}$$

$$\begin{aligned}K_t &\equiv \sum_{k=0}^{\infty}(\theta\beta)^k E_t\biggl\{\frac{1}{1-\tau_{t+k}}\frac{\mathcal{M}}{1-\alpha}\biggl(\frac{Y_{t+k}}{A_{t+k}}\biggr)^{\frac{1+\varphi}{1-\alpha}}\biggl(\frac{P_{t+k}}{P_t}\biggr)^{\frac{e(1+\varphi)}{1-\alpha}}\biggr\}\\ F_t &\equiv \sum_{k=0}^{\infty}(\theta\beta)^k E_t\biggl\{U_{c,t+k}\biggl(\frac{P_{t+k}}{P_t}\biggr)^{e-1}Y_{t+k}\biggr\}\\ \omega &\equiv \frac{1+\varphi}{1-\alpha}-1\end{aligned}$$

A second order Taylor series for the left hand-side of (3.54) takes the form:

$$-log\biggl(\frac{1-\theta\Pi_t^{e-1}}{1-\theta}\biggr)=\frac{\theta}{1-\theta}(e-1)\biggl\{\pi_t+\frac{1}{2}\frac{e-1}{1-\theta}\pi_t^2+\mathcal{O}(||\xi||^3)\biggr\},$$

whereas second order approximations for $log(K_t)$ and $log(F_t)$ imply:

$$\begin{aligned}\widehat{K_t}+\frac{1}{2}\widehat{K_t}^2+\mathcal{O}(||\xi||^3) &= (1-\theta\beta)E_t\sum_{T=t}^{\infty}(\theta\beta)^{T-t}\Bigl[\widehat{k}_{t,T}+\frac{1}{2}\widehat{k}_{t,T}^2\Bigr]+\mathcal{O}(||\xi||^3)\\ \widehat{F}_t+\frac{1}{2}\widehat{F}_t^2+\mathcal{O}(||\xi||^3) &= (1-\theta\beta)E_t\sum_{T=t}^{\infty}(\theta\beta)^{T-t}\Bigl[\widehat{f}_{t,T}+\frac{1}{2}\widehat{f}_{t,T}^2\Bigr]+\mathcal{O}(||\xi||^3)\end{aligned}$$

where:

$$\widehat{k}_{t,T} \equiv \widehat{k}_T + e(1+\omega)\sum_{s=t+1}^{T}\pi_s, \quad k_T \equiv log\Bigg(\frac{1}{1-\tau_T}\mathcal{M}\frac{1}{1-\alpha}\Big(\frac{Y_T}{A_T}\Big)^{\frac{1+\varphi}{1-\alpha}}\Bigg)$$

$$\widehat{f}_{t,T} \equiv \widehat{f}_T + (e-1)\sum_{s=t+1}^{T}\pi_s, \quad f_T \equiv log(U_{Y,T}Y_T) = log(Z_T Y_T^{1-\sigma})$$

A second order approximation of $log(1-\tau_t)$ gives:

$$\begin{aligned} log(1-\tau_t) &= log(1-\tau) + \frac{(-1)\tau}{1-\tau}\frac{\Delta\tau_t^w}{\tau} + \frac{1}{2}(-1)\frac{-(\tau)^2}{(1-\tau)^2}(-1)\Bigg(\frac{\Delta\tau_t^w}{\tau}\Bigg)^2 + \\ &+ \mathcal{O}(||\xi||^3) \\ &= log(1-\tau) - \frac{\tau}{1-\tau}\widehat{\tau}_t^w - \frac{1}{2}\frac{\tau}{(1-\tau)^2}(\widehat{\tau}_t^w)^2 + \mathcal{O}(||\xi||^3), \end{aligned}$$

Hence,

$$\widehat{k}_T = \frac{1+\varphi}{1-\alpha}\Bigg(\widehat{y}_T - a_T\Bigg) + \frac{\tau}{1-\tau}\widehat{\tau}_t + \frac{1}{2}\frac{\tau}{(1-\tau)^2}(\widehat{\tau}_t)^2 + \mathcal{O}(||\xi||^3) \tag{3.55}$$

$$\widehat{f}_T = z_T + (1-\sigma)\widehat{y}_T \tag{3.56}$$

As shown by Benigno and Woodford (2004), the second order approximations of the right and left hand sides of the aggregate supply relation (3.54) yield the following relation:

$$\begin{aligned} \pi_t + \frac{1}{2}\frac{e-1}{1-\theta}\pi_t^2 + \frac{1}{2}(1-\theta\beta)\pi_t\mathcal{Z}_t &= \frac{1-\theta}{\theta}\frac{1-\theta\beta}{1+\omega e}\Bigg[(\widehat{k}_t - \widehat{f}_t) + \frac{1}{2}(\widehat{k}_t^2 - \widehat{f}_t^2)\Bigg] + \\ &+ \beta E_t\pi_{t+1} + \beta E_t\Bigg[\frac{1}{2}\frac{e-1}{1-\theta}\pi_{t+1}^2\Bigg] + \beta\frac{1}{2}(1-\theta\beta)E_t\pi_{t+1}\mathcal{Z}_{t+1} \\ &+ \beta\frac{1}{2}e(1+\omega)E_t\pi_{t+1}^2 + \mathcal{O}(||\xi||^3) \end{aligned} \tag{3.57}$$

$$\text{where} \quad \mathcal{Z}_t \equiv E_t\sum_{T=t}^{\infty}(\theta\beta)^{T-t}\Big[\widehat{k}_{t,T} + \widehat{f}_{t,T}\Big]$$

We define $V_t \equiv \pi_t + \frac{1}{2}\Bigg(\frac{e-1}{1-\theta} + e(1+\omega)\Bigg)\pi_t^2 + \frac{1}{2}(1-\theta\beta)\pi_t\mathcal{Z}_t$ and use relations

(3.55) and (3.56) to compute:

$$\widehat{k}_t - \widehat{f}_t = \left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)\widehat{y}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t + \frac{1}{2}\frac{\tau}{(1-\tau)^2}(\widehat{\tau}_t)^2 + t.i.p. + \mathcal{O}(||\xi||^3)$$

$$\frac{1}{2}\left(\widehat{k}_t^2 - \widehat{f}_t^2\right) = \frac{1}{2}\left[\left(\frac{1+\varphi}{1-\alpha}\right)^2 - (1-\sigma)^2\right]\widehat{y}_t^2 + \frac{1}{2}\left(\frac{\tau}{1-\tau}\right)^2(\widehat{\tau}_t)^2 + \frac{(1+\varphi)}{1-\alpha}\frac{\tau}{1-\tau}\widehat{y}_t\widehat{\tau}_t$$
$$- \left(\frac{1+\varphi}{1-\alpha}\right)^2\widehat{y}_t a_t - (1-\sigma)\widehat{y}_t z_t - \frac{1+\varphi}{1-\alpha}\frac{\tau}{1-\tau}\widehat{\tau}_t a_t + t.i.p. + \mathcal{O}(||\xi||^3)$$

in order to write (3.57) as:

$$V_t = \kappa\left\{c_x' x_t + \frac{1}{2}x_t' C_x x_t + x_t' C_\xi \xi_t + \frac{1}{2}c_\pi \pi_t^2\right\} + \beta E_t V_{t+1} + \mathcal{O}(||\xi||^3) + t.i.p.$$

where: $\kappa \equiv \nu_k\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right), \quad \nu_k \equiv \frac{1-\theta}{\theta}\frac{1-\theta\beta}{1+\omega e}$

$$c_x' \equiv \left[\frac{\tau}{1-\tau}\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1} \quad 1\right]$$

$x_t \equiv [\widehat{\tau}_t \;\; \widehat{y}_t]'$

$\xi_t \equiv [z_t \;\; a_t]'$

$$C_x \equiv \left[\frac{\tau(1+\tau)}{(1-\tau)^2}\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1} \quad \frac{(1+\varphi)}{1-\alpha}\frac{\tau}{1-\tau}\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1};\right.$$
$$\left.\frac{(1+\varphi)}{1-\alpha}\frac{\tau}{1-\tau}\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1} \quad \left[\left(\frac{1+\varphi}{1-\alpha}\right)^2 - (1-\sigma)^2\right]\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1}\right]$$

$$C_\xi \equiv \left[0, \;\; -\frac{\tau}{1-\tau}\frac{1+\varphi}{1-\alpha}\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1};\right.$$
$$\left. -(1-\sigma)\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1} - \left(\frac{1+\varphi}{1-\alpha}\right)^2\left(\sigma + \frac{\alpha+\varphi}{1-\alpha}\right)^{-1}\right]$$

$c_\pi \equiv e(1+\omega)\kappa^{-1}$

We can integrate the equation above forward from time t to obtain:

$$\kappa^{-1}V_t = E_t\sum_{k=0}^{\infty}\beta^k\left\{c_x' x_{t+k} + \frac{1}{2}x_{t+k}' C_x x_{t+k} + x_{t+k}' C_\xi \xi_{t+k} + \frac{1}{2}c_\pi \pi_{t+k}^2\right\} \quad (3.58)$$
$$+ \mathcal{O}(||\xi||^3) + t.i.p. \quad (3.59)$$

Next, we determine a second order approximation to the intertemporal government solvency condition. The flow budget constraint of the government (3.9) implies:

$$\Pi_t^{-1}\frac{B_{t-1}}{P_{t-1}} = Q_t\frac{B_t}{P_t} + s_t, \quad s_t \equiv \tau_t\frac{W_t}{P_t}L_t - \mathcal{T}$$

where s_t is the government primary (real) surplus. This constraint can be iterated forward to get (after imposing the government public solvency condition and using the households' consumption/saving equation):

$$W_t = \sum_{k=0}^{\infty}\beta^k E_t\{U_{c,t+k}s_{t+k}\}, \quad W_t \equiv \Pi_t^{-1}\frac{B_{t-1}}{P_{t-1}}U_{c,t} \tag{3.60}$$

A second order approximation of $U_{c,t}s_t = Z_tC_t^{-\sigma}s_t$ gives:

$$\begin{aligned}U_{c,t}s_t &= U_cs + Zs(-\sigma)C^{-\sigma-1}\Delta C_t + ZC^{-\sigma}\Delta s_t +\\ &+ \frac{1}{2}Zs(-\sigma)(-\sigma-1)C^{-\sigma-2}(\Delta C_t)^2 + s(-\sigma)C^{-\sigma-1}\Delta Z_t\Delta C_t +\\ &+ C^{-\sigma}\Delta Z_t\Delta s_t + Z(-\sigma)C^{-\sigma-1}\Delta C_t\Delta s_t + \mathcal{O}(||\xi||^3) + t.i.p.\\ &= C^{-\sigma}s\left[1 - \sigma\widehat{y}_t + \frac{1}{2}\sigma^2\widehat{y}_t^{\,2} + \frac{\Delta s_T}{s} - \sigma\widehat{y}_t\widehat{z}_t + \left(\Delta Z_t - \sigma\frac{\Delta Y_t}{Y}\right)\frac{\Delta s_t}{s}\right]\\ &+ \mathcal{O}(||\xi||^3) + t.i.p.\end{aligned}$$

Using the labor supply relation (3.4) and the goods market clearing condition $Y_t = C_t$ we can write the government primary surplus as:

$$s_t = \frac{1}{(\tau_t^{-1}-1)}Y_t^{\sigma}L_t^{\varphi+1} - \mathcal{T},$$

which can be approximated up to second order by:

$$\begin{aligned}s_t =& s + \frac{\tau Y^\sigma L^{1+\varphi}}{(1-\tau)^2}\frac{\Delta\tau_t}{\tau} + \frac{\sigma Y^\sigma L^{1+\varphi}}{\tau^{-1}-1}\frac{\Delta Y_t}{Y} + \frac{(1+\varphi)Y^\sigma L^{\varphi+1}}{\tau^{-1}-1}\frac{\Delta L_t}{L} +\\ &+ Y^\sigma L^{1+\varphi}(1-\tau)^{-3}\tau^2\left(\frac{\Delta\tau_t}{\tau}\right)^2 + \frac{1}{2}\frac{\sigma(\sigma-1)Y^\sigma L^{1+\varphi}}{\tau^{-1}-1}\left(\frac{\Delta Y_t}{Y}\right)^2 +\\ &+ \frac{1}{2}\frac{(1+\varphi)\varphi Y^\sigma L^{1+\varphi}}{\tau^{-1}-1}\left(\frac{\Delta L_t}{L}\right)^2 + \frac{\tau\sigma Y^\sigma L^{1+\varphi}}{(1-\tau)^2}\frac{\Delta\tau_t}{\tau}\frac{\Delta Y_t}{Y} +\\ &+ \frac{\tau(1+\varphi)Y^\sigma L^{1+\varphi}}{(1-\tau)^2}\frac{\Delta\tau_t}{\tau}\frac{\Delta L_t}{L} + \frac{(1+\varphi)\sigma Y^\sigma L^{1+\varphi}}{\tau^{-1}-1}\frac{\Delta Y_t}{Y}\frac{\Delta L_t}{L} +\\ &+ \mathcal{O}(||\xi||^3)\end{aligned}$$

$$\text{with } \omega_\tau \equiv s^{-1}\frac{\tau Y^\sigma L^{1+\varphi}}{(1-\tau)^2} = (s/Y)^{-1}\frac{\tau}{1-\tau}\mathcal{M}^{-1}(1-\alpha)$$

$$\omega_{\tau\tau} \equiv s^{-1}\left[\frac{1}{2}\frac{\tau Y^\sigma L^{1+\varphi}}{(1-\tau)^2} + Y^\sigma L^{1+\varphi}(1-\tau)^{-3}\tau^2\right]$$

$$= (s/Y)^{-1}\frac{1-\alpha}{\mathcal{M}}\frac{\tau}{1-\tau}\left(\frac{1}{2}+\frac{\tau}{1-\tau}\right)$$

$$\omega_y \equiv s^{-1}\left[\frac{\sigma Y^\sigma L^{1+\varphi}}{\tau^{-1}-1} + \frac{(1+\varphi)Y^\sigma L^{\varphi+1}}{(\tau^{-1}-1)(1-\alpha)}\right]$$

$$= (s/Y)^{-1}\frac{(1-\tau)(1-\alpha)}{\mathcal{M}(\tau^{-1}-1)}\left(\sigma+\frac{1+\varphi}{1-\alpha}\right)$$

$$\omega_{yy} \equiv s^{-1}\left[\frac{1}{2}\frac{\sigma Y^\sigma L^{1+\varphi}}{\tau^{-1}-1} + \frac{1}{2}\frac{(1+\varphi)^2 Y^\sigma L^{\varphi+1}}{(\tau^{-1}-1)(1-\alpha)^2} + \frac{1}{2}\frac{\sigma(\sigma-1)Y^\sigma L^{1+\varphi}}{\tau^{-1}-1} + \right.$$

$$\left. + \frac{(1+\varphi)\sigma Y^\sigma L^{1+\varphi}}{(\tau^{-1}-1)(1-\alpha)}\right]$$

$$= (s/Y)^{-1}\frac{(1-\tau)(1-\alpha)}{(\tau^{-1}-1)\mathcal{M}}\left[\frac{1}{2}\frac{(1+\varphi)^2}{(1-\alpha)^2} + \frac{1}{2}\sigma^2 + \frac{\sigma(1+\varphi)}{1-\alpha}\right]$$

$$\omega_\pi \equiv s^{-1}\frac{(1+\varphi)Y^\sigma L^{\varphi+1}}{(\tau^{-1}-1)(1-\alpha)}\frac{\varepsilon}{2\Theta}$$

$$= (s/Y)^{-1}\frac{(1+\varphi)(1-\tau)}{(\tau^{-1}-1)\mathcal{M}}\frac{\varepsilon}{2\Theta}, \quad \Theta \equiv \frac{1-\alpha}{1-\alpha+\alpha\epsilon}$$

$$\omega_{\tau y} \equiv s^{-1}\left[\frac{\tau\sigma Y^\sigma L^{1+\varphi}}{(1-\tau)^2} + \frac{\tau(1+\varphi)Y^\sigma L^{1+\varphi}}{(1-\tau)^2(1-\alpha)}\right] = (s/Y)^{-1}\frac{\tau}{1-\tau}\frac{1-\alpha}{\mathcal{M}}\left(\sigma+\frac{1+\varphi}{1-\alpha}\right)$$

$$\omega_{ya} \equiv -s^{-1}\left[\frac{(1+\varphi)Y^\sigma L^{\varphi+1}}{(\tau^{-1}-1)(1-\alpha)^2} + \frac{(1+\varphi)\varphi Y^\sigma L^{1+\varphi}}{(\tau^{-1}-1)(1-\alpha)^2} + \frac{(1+\varphi)\sigma Y^\sigma L^{1+\varphi}}{(\tau^{-1}-1)(1-\alpha)}\right]$$

$$= -(s/Y)^{-1}\frac{(1+\varphi)(1-\tau)}{(\tau^{-1}-1)\mathcal{M}}\left(\sigma+\frac{1+\varphi}{1-\alpha}\right)$$

$$\omega_{\tau a} \equiv -s^{-1}\frac{\tau(1+\varphi)Y^\sigma L^{1+\varphi}}{(1-\tau)^2(1-\alpha)}$$

$$= -(s/Y)^{-1}\frac{\tau}{1-\tau}\frac{1+\varphi}{\mathcal{M}}$$

$$\omega_a \equiv -s^{-1}\frac{(1+\varphi)Y^\sigma L^{1+\varphi}}{(1-\alpha)(\tau^{-1}-1)} = -(s/Y)^{-1}\frac{(1+\varphi)(1-\tau)}{(\tau^{-1}-1)\mathcal{M}}$$

$$\text{So,}\quad U_{c,T}s_T = C^{-\sigma}s\Big[\omega_\tau\widehat{\tau}_t + (\omega_y - \sigma)\widehat{y}_t + \frac{1}{2}(\sigma^2 + 2\omega_{yy} - 2\sigma\omega_y)\widehat{y}_t^{\,2} + $$
$$+ \frac{1}{2}2\omega_{\tau\tau}\widehat{\tau}_t^2 + \frac{1}{2}2(\omega_{\tau y} - \sigma\omega_\tau)\widehat{\tau}_t\widehat{y}_t + (\omega_{ya} - \sigma\omega_a)\widehat{y}_t a_t +$$
$$+ (\omega_y - \sigma)\widehat{y}_t z_t + \omega_{\tau a}\widehat{\tau}_t a_t + \omega_\tau\widehat{\tau}_t z_t + \omega_\pi var_i\{p_t(i)\}\Big] +$$
$$+ \mathcal{O}(||\xi||^3) + t.i.p.$$

A second order approximation of W_t in (3.60) thus yields:

$$\Delta W_t = \sum_{k=0}^{\infty}\beta^k E_t\Big\{C^{-\sigma}s\Big[\omega_\tau\widehat{\tau}_{t+k} + (\omega_y - \sigma)\widehat{y}_{t+k} + (\frac{1}{2}\sigma^2 + \omega_{yy} - \sigma\omega_y)\widehat{y}_{t+k}^2 +$$
$$+ \omega_{\tau\tau}\widehat{\tau}_{t+k}^2 + (\omega_{\tau y} - \sigma\omega_\tau)\widehat{\tau}_{t+k}\widehat{y}_{t+k} + (\omega_{ya} - \sigma\omega_a)\widehat{y}_{t+k}a_{t+k} +$$
$$+ (\omega_y - \sigma)\widehat{y}_{t+k}z_{t+k} + \omega_{\tau a}\widehat{\tau}_{t+k}a_{t+k} + \omega_\tau\widehat{\tau}_{t+k}z_{t+k} + \omega_\pi var_i\{p_{t+k}(i)\}\Big]\Big\}$$
$$+ \mathcal{O}(||\xi||^3) + t.i.p.,$$

which can be further written as:

$$\Delta W_t = C^{-\sigma}s\Big[\omega_\tau\widehat{\tau}_t + (\omega_y - \sigma)\widehat{y}_t + \frac{1}{2}(\sigma^2 + 2\omega_{yy} - 2\sigma\omega_y)\widehat{y}_t^{\,2} + \frac{1}{2}2\omega_{\tau\tau}\widehat{\tau}_t^2 +$$
$$\frac{1}{2}2(\omega_{\tau y} - \sigma\omega_\tau)\widehat{\tau}_t\widehat{y}_t + (\omega_{ya} - \sigma\omega_a)\widehat{y}_t a_t + (\omega_y - \sigma)\widehat{y}_t z_t + \omega_{\tau a}\widehat{\tau}_t a_t +$$
$$+ \omega_\tau\widehat{\tau}_t z_t + \omega_\pi var_i\{p_t(i)\}\Big] + \beta E_t\Delta W_{t+1} + \mathcal{O}(||\xi||^3) + t.i.p.,$$

and, dividing by $W = \frac{sU_c}{(1-\beta)}$ and using the notation $\widetilde{W}_t \equiv \frac{W_t - W}{W}$, as:

$$\widetilde{W}_t = (1-\beta)\Big[b_x' x_t + \frac{1}{2}x_t' B_x x_t + x_t' B_\xi \xi_t + \omega_\pi var_i\{p_t(i)\}\Big]$$
$$+ \beta E_t\big\{\widetilde{W}_{t+1}\big\} + \mathcal{O}(||\xi||^3) + t.i.p.$$

$$\text{with}\quad b_x' \equiv [\omega_\tau \quad (\omega_y - \sigma)]$$
$$B_x \equiv [2\omega_{\tau\tau} \quad (\omega_{\tau y} - \sigma\omega_\tau); (\omega_{\tau y} - \sigma\omega_\tau) \quad (\sigma^2 + 2\omega_{yy} - 2\sigma\omega_y)]$$
$$B_\xi \equiv [\omega_\tau \quad \omega_{\tau a;}(\omega_y - \sigma) \quad (\omega_{ya} - \sigma\omega_a)]$$

Integrating this equation forward and using:

$$\sum_{t=0}^{\infty}\beta^t var_i\{p_t(i)\} = \frac{\theta}{(1-\beta\theta)(1-\theta)}\sum_{t=0}^{\infty}\beta^t\pi_t^2$$

from Woodford (2003), Chapter 6, we obtain:

$$
\begin{aligned}
(1-\beta)^{-1}\widetilde{W}_t =& E_t \sum_{k=0}^{\infty} \beta^k \left[b_x' x_t + \frac{1}{2} x_t' B_x x_t + x_t' B_\xi \xi_t + \frac{1}{2} b_\pi \pi_t^2 \right] + \\
&+ \mathcal{O}(||\xi||^3) + t.i.p. \qquad (3.61)
\end{aligned}
$$

with $b_\pi \equiv 2\omega_\pi \frac{\theta}{(1-\beta\theta)(1-\theta)}$. We can now express the linear term in (3.52) in terms of quadratic terms by combining (3.58) with (3.61). We do so by finding ν_1 and ν_2 such that:

$$\nu_1 b_x' + \nu_2 c_x' \equiv [0 \quad \Phi],$$

$$
\begin{aligned}
\text{namely} \quad \nu_2 &\equiv \Phi \left[1 - \frac{\tau(\omega_y - \sigma)(1-\alpha)}{(1-\tau)\omega_\tau[\sigma(1-\alpha)+\alpha+\varphi]} \right]^{-1} \\
\nu_1 &\equiv -\frac{\tau(1-\alpha)}{(1-\tau)\omega_\tau[\sigma(1-\alpha)+\alpha+\varphi]} \nu^2
\end{aligned}
$$

These two values allow to write:

$$
\begin{aligned}
E_t \sum_{k=0}^{\infty} \beta^k \Phi \widehat{y}_{t+k} &= E_t \sum_{k=0}^{\infty} \beta^k [\nu_1 b_x' + \nu_2 c_x'] x_{t+k} \\
&= -E_t \sum_{k=0}^{\infty} \beta^k \left[\frac{1}{2} x_{t+k}' D_x x_{t+k} + x_{t+k}' D_\xi \xi_{t+k} + \frac{1}{2} d_\pi \pi_{t+k}^2 \} \right] \\
&+ \nu_1 (1-\beta)^{-1} \widetilde{W}_t + \nu_2 \kappa^{-1} V_t + \mathcal{O}(||\xi||^3) + t.i.p.
\end{aligned}
$$

$$
\begin{aligned}
\text{with} \quad D_x &\equiv \nu_1 B_x + \nu_2 C_x \\
D_\xi &\equiv \nu_1 B_\xi + \nu_2 C_\xi \\
d_\pi &\equiv \nu_1 b_\pi + \nu_2 c_\pi
\end{aligned}
$$

Replacing this expression in (3.53), we get the following quadratic expression for welfare:

$$\mathcal{WL} = E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2} x_t' Q_x x_t + x_t' Q_\xi \xi_t + \frac{1}{2} q_\pi \pi_t^2 \right) + T_0 + \mathcal{O}(||\xi||^3) + t.i.p.$$

$$
\begin{aligned}
\text{with} \quad Q_x &\equiv \left[D_x^{11} \quad D_x^{12}; D_x^{21} \quad D_x^{22} + u_{yy} \right] \\
Q_\xi &\equiv \left[D_\xi^{11} \quad D_\xi^{12}; D_\xi^{21} - \Phi \quad D_\xi^{22} - \frac{(1-\Phi)(1+\varphi)}{1-\alpha} \right] \\
q_\pi &\equiv d_\pi + u_\pi \\
T_0 &\equiv -\nu_1 (1-\beta)^{-1} \widetilde{W}_0 - \nu_2 \kappa^{-1} V_0
\end{aligned}
$$

which can be rephrased in terms of deviations from a target level of output as:

$$\mathcal{WL} = E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2} Q_x^{22} \widetilde{y}_t^2 + \frac{1}{2} Q_x^{11} \widehat{\tau}_t^2 + Q_x^{12} \widetilde{y}_t \widehat{\tau}_t + \xi_t^{\tau} \widehat{\tau}_t + \frac{1}{2} q_{\pi} \pi_t^2 \right) + \\ + T_0 + \mathcal{O}(||\xi||^3) + t.i.p.$$

where $\widetilde{y}_t \equiv \widehat{y}_t - \widehat{y}_t^*$, $\widehat{y}_t^* \equiv -\frac{Q_\xi^{21}}{Q_x^{22}} z_t - \frac{Q_\xi^{22}}{Q_x^{22}} a_t$ and $\xi_t^{\tau} \equiv \left(Q_\xi^{11} - \frac{Q_x^{12} Q_\xi^{21}}{Q_x^{22}} \right) z_t + \left(Q_\xi^{12} - \frac{Q_x^{12} Q_\xi^{22}}{Q_x^{22}} \right) a_t$.

Following Eggertsson and Woodford (2004), we will rank policies in terms of the implied value of the discounted quadratic loss function:

$$\mathcal{WL} = E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2} Q_x^{22} \widetilde{y}_t^2 + \frac{1}{2} Q_x^{11} \widehat{\tau}_t^2 + Q_x^{12} \widetilde{y}_t \widehat{\tau}_t + \xi_t^{\tau} \widehat{\tau}_t + \frac{1}{2} q_{\pi} \pi_t^2 \right)$$

Because this loss function is purely quadratic it is possible to evaluate it to second order using only a first order approximation to the equilibrium evolution of inflation and output under a given policy.

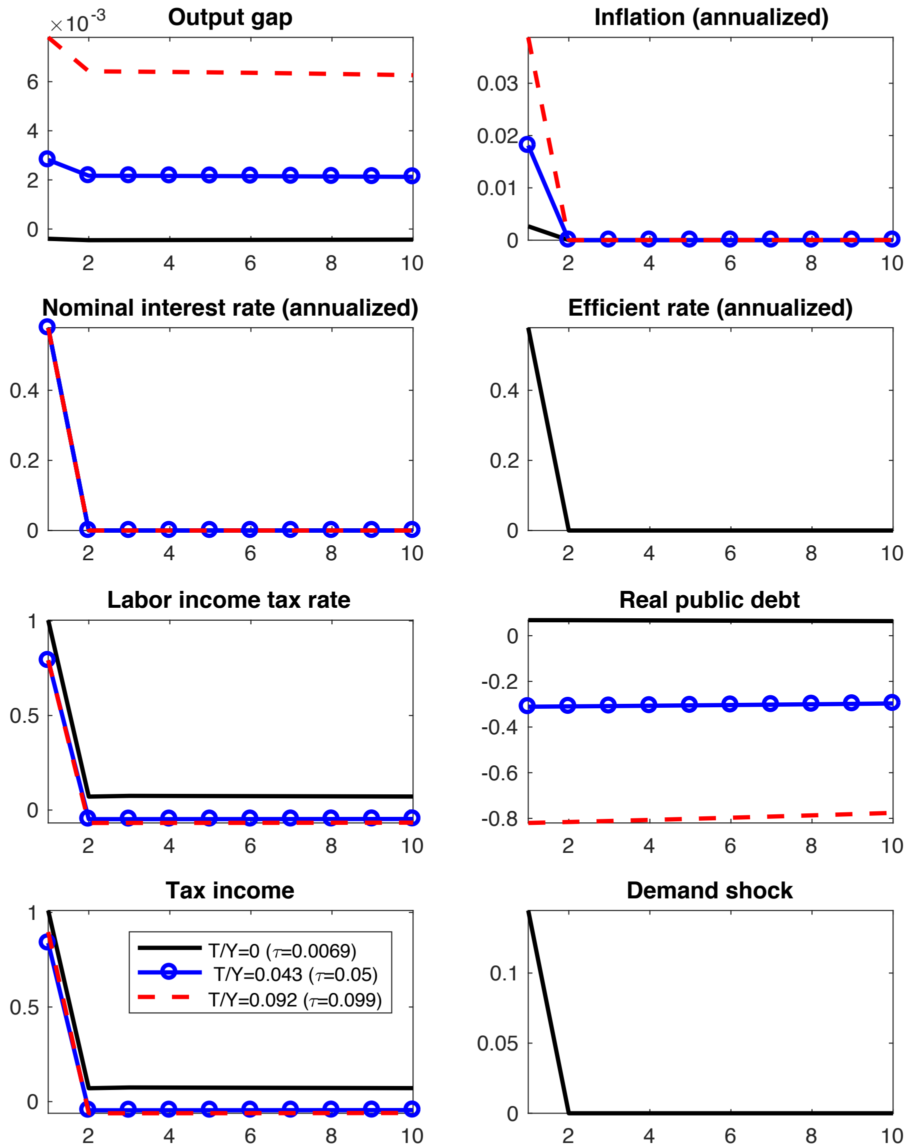


Figure 3.18: Dynamic responses to a one period positive demand preference shock under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$, no ZLB)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

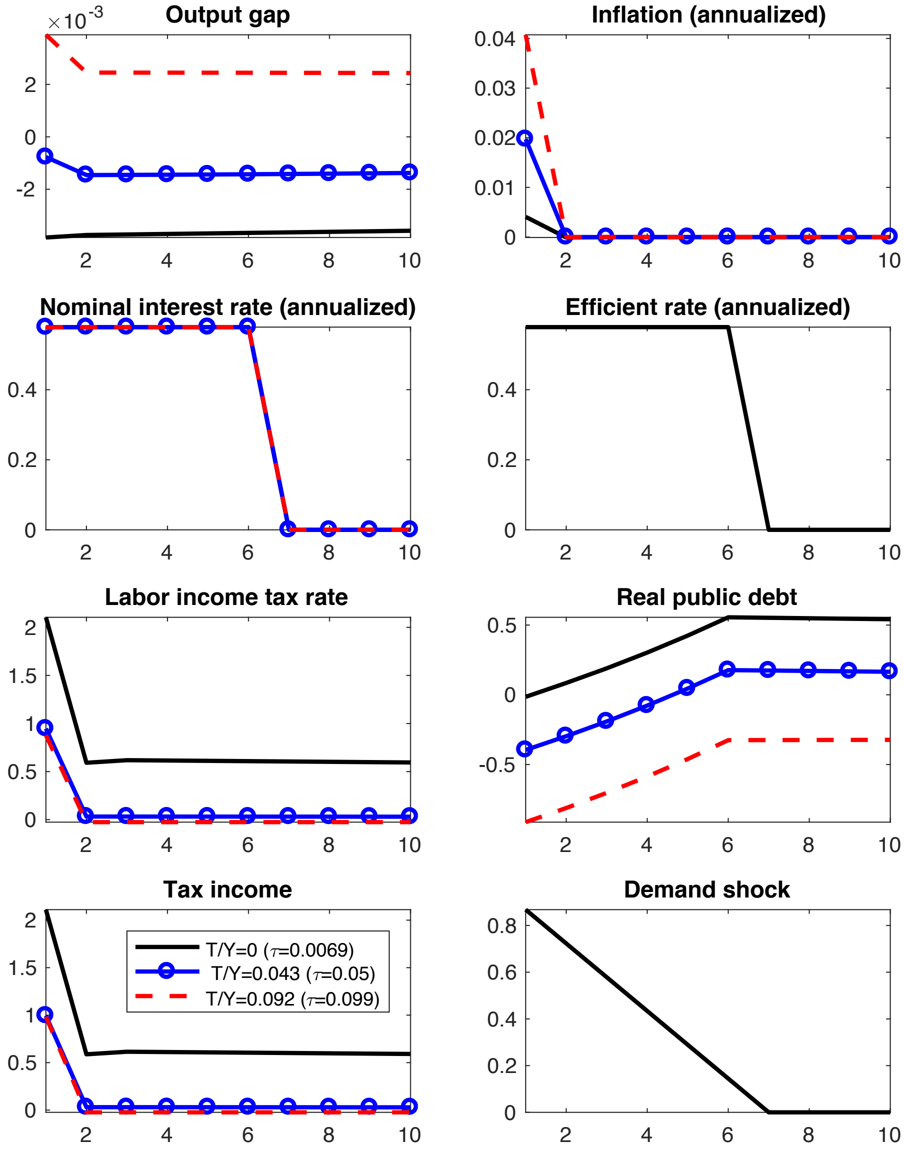


Figure 3.19: Dynamic responses to a six period positive demand preference shock under optimal policy (small steady-state distorsions with $\bar{b} = 0.1$, no ZLB)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

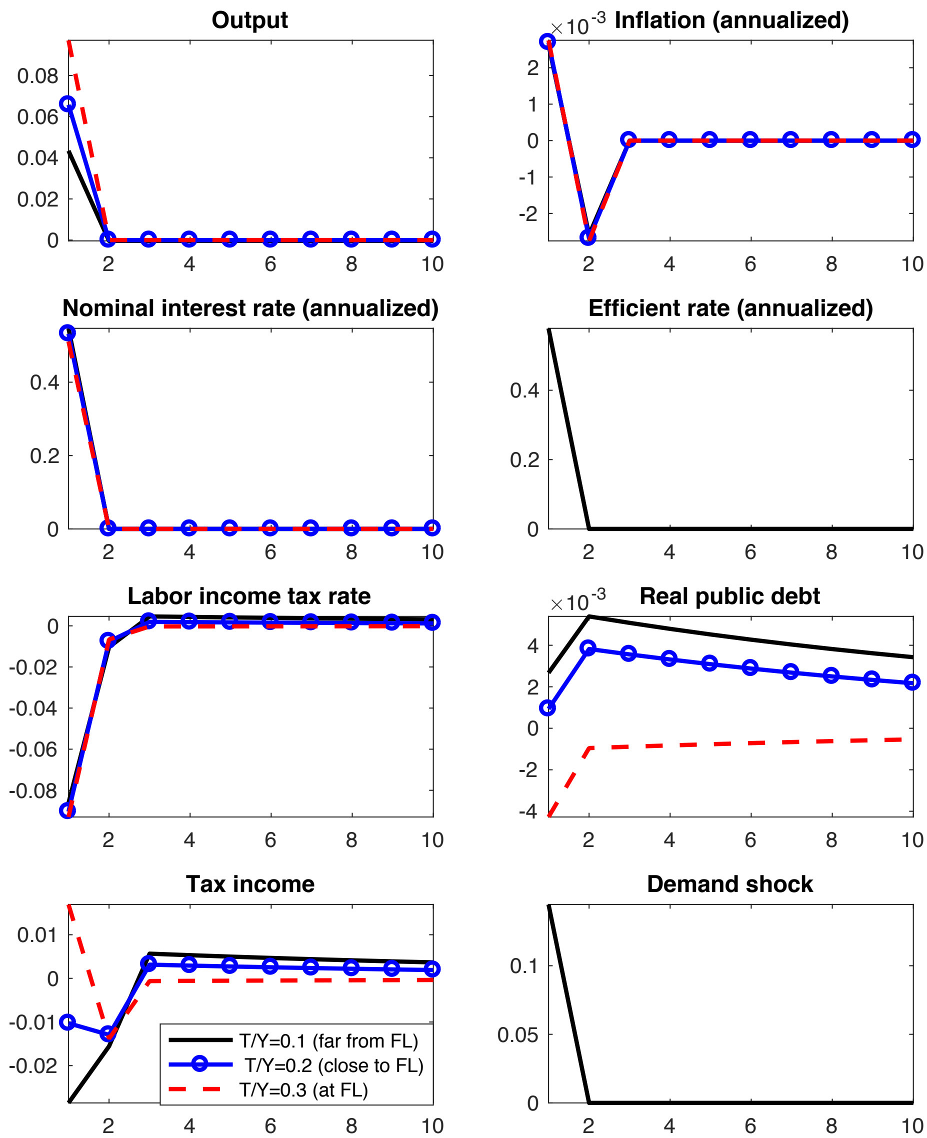


Figure 3.20: Dynamic responses to a transitory positive demand shock under optimal policy (large steady-state distorsions)

Note: Y-axis: % deviation from steady-state. X-axis: quarters

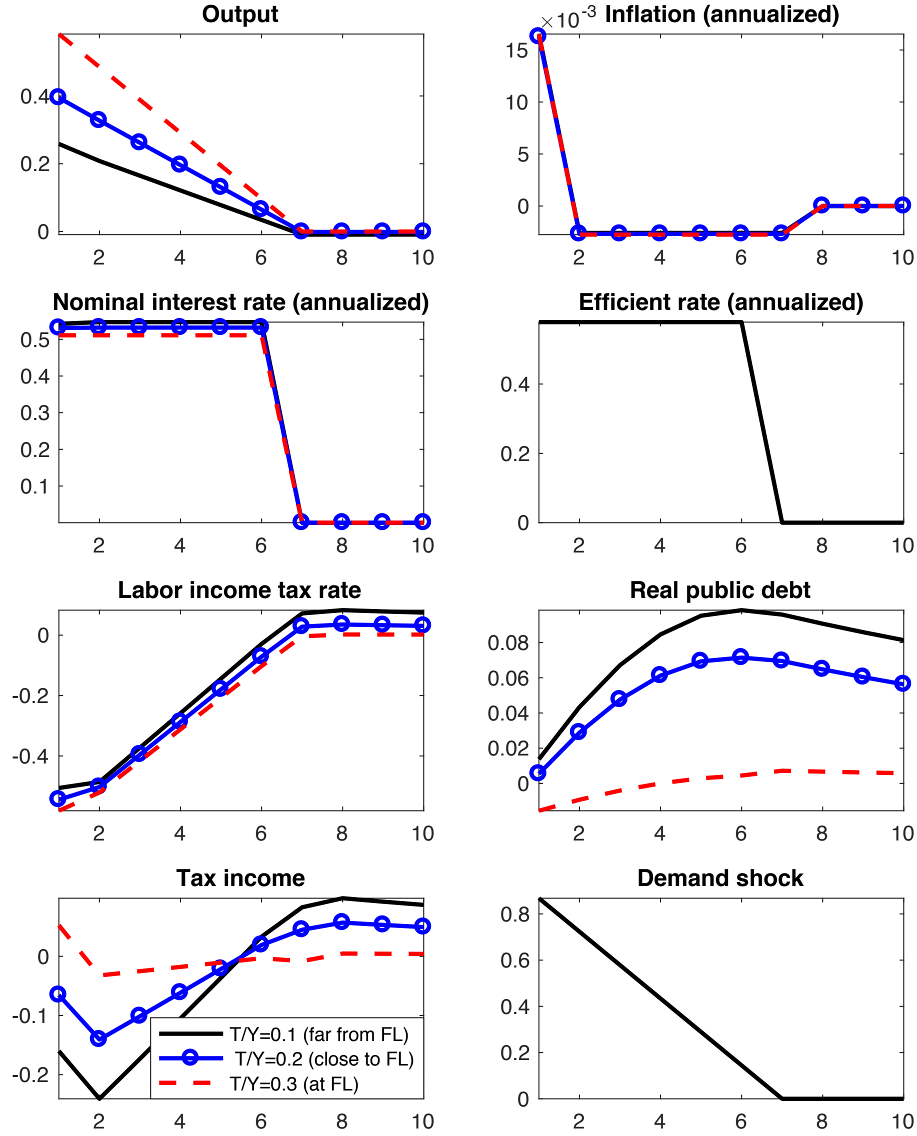


Figure 3.21: Dynamic responses to a six period positive demand shock under optimal policy (large steady-state distorsions)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

3.7.7 Case of zero steady-state debt

Constraints optimization for $B/P = 0$ and small steady-state distorsions

$$\text{DIS:} \quad \widehat{\widetilde{y}}_t = E_t\{\widehat{\widetilde{y}}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\} - \widehat{r}^e_t\Big)$$

$$\text{NKPC:} \; \pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\Big(\sigma + \frac{\alpha+\varphi}{1-\alpha}\Big)\widehat{\widetilde{y}}_t + \lambda\frac{\tau}{1-\tau}\widehat{\tau}_t$$

$$\widehat{b}^y_t = \beta^{-1}\widehat{b}^y_{t-1} - \bar{t}\beta^{-1}\Big[\frac{1}{1-\tau}\widehat{\tau}_t + \Big(\sigma + \frac{1+\varphi}{1-\alpha}\Big)(\widetilde{y}_t + \widehat{y}^e_t) - \frac{1+\varphi}{1-\alpha}a_t\Big]$$

$$\lim_{T\to\infty}\beta^{T+1}E_t\{\widehat{b}^y_T\} = 0$$

$$\Big(\widehat{i}_t + \rho \geq 0\Big)$$

Derivations:

$$\begin{aligned}
\widehat{y}_t &= E_t\{\widehat{y}_{t+1}\} - \frac{1}{\sigma}(\widehat{i}_t - E_t\{\pi_{t+1}\}) + \frac{1}{\sigma}(1-\rho_z)z_t \\
\widehat{y}^e_t &= E_t\{\widehat{y}^e_{t+1}\} - \frac{1}{\sigma}\widehat{r}^e_t + \frac{1}{\sigma}(1-\rho_z)z_t \\
\Rightarrow DIS: \; \widehat{\widetilde{y}}_t &= E_t\{\widehat{\widetilde{y}}_{t+1}\} - \frac{1}{\sigma}(\widehat{i}_t - E_t\{\pi_{t+1}\} - \widehat{r}^e_t)
\end{aligned}$$

$$\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\widehat{mc}_t$$

$$\begin{aligned}
0 = \widehat{mc}^e_t &= \widehat{\omega}^e_t + \widehat{l}^e_t - \widehat{y}^e_t \\
&= (\sigma\widehat{y}^e_t + \varphi\widehat{l}^e_t) + \widehat{l}^e_t - \widehat{y}^e_t \\
&= (\sigma-1)\widehat{y}^e_t + (\varphi+1)\widehat{l}^e_t \\
&= (\sigma-1)\widehat{y}^e_t + \frac{\varphi+1}{1-\alpha}(\widehat{y}^e_t - a_t) \\
&= \Big(\sigma - 1 + \frac{\varphi+1}{1-\alpha}\Big)\widehat{y}^e_t - \frac{\varphi+1}{1-\alpha}a_t \\
&= \Big(\sigma + \frac{\varphi+\alpha}{1-\alpha}\Big)\widehat{y}^e_t - \frac{1+\varphi}{1-\alpha}a_t(1)
\end{aligned}$$

$$
\begin{aligned}
\widehat{mc}_t &= \widehat{\omega}_t + \widehat{l}_t - \widehat{y}_t \\
&= (\sigma \widehat{y}_t + \varphi \widehat{l}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t) + \widehat{l}_t - \widehat{y}_t \\
&= \Big(\sigma + \frac{\varphi+\alpha}{1-\alpha}\Big)\widehat{y}_t - \frac{1+\varphi}{1-\alpha}a_t + \frac{\tau}{1-\tau}\widehat{\tau}_t (2) \\
(2)-(1) \Rightarrow \widehat{mc}_t &= \Big(\sigma + \frac{\varphi+\alpha}{1-\alpha}\Big)\widehat{\widetilde{y}}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t
\end{aligned}
$$

$$
\Rightarrow NKPC : \pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\Big(\sigma + \frac{\varphi+\alpha}{1-\alpha}\Big)\widehat{\widetilde{y}}_t + \lambda\frac{\tau}{1-\tau}\widehat{\tau}_t
$$

$$
\begin{aligned}
Q_t\frac{B_t}{P_t} &= \frac{B_{t-1}}{P_{t-1}}\Pi_t^{-1} + (\mathcal{T} - \tau_t\frac{W_t}{P_t}L_t) \\
\beta e^{\widehat{q}_t}\frac{B_t}{P_t} &= \frac{B_{t-1}}{P_{t-1}}e^{-\pi_t} + (\mathcal{T} - \tau\frac{W}{P}Le^{\widehat{\tau}_t+\widehat{\omega}_t+\widehat{l}_t}) \\
\beta e^{\widehat{q}_t}\frac{B_t/P_t}{Y} &= \frac{B_{t-1}/P_{t-1}}{Y}e^{-\pi_t} + \Big(\frac{\mathcal{T}}{Y} - \tau\frac{W}{P}\frac{L}{Y}e^{\widehat{\tau}_t+\widehat{\omega}_t+\widehat{l}_t}\Big)
\end{aligned}
$$

Up to a first order approximation,

$$
\beta\bar{b}\widehat{q}_t + \beta\frac{\Delta B_t/P_t}{Y} = -\bar{b}\pi_t + \Delta\frac{B_{t-1}/P_{t-1}}{Y} - \tau\frac{W}{P}\frac{L}{Y}(\widehat{\tau}_t + \widehat{\omega}_t + \widehat{l}_t)
$$

with $\bar{b} \equiv \frac{B/P}{Y}$. In steady-state with $B/P = 0$, $\tau\frac{W}{P}\frac{L}{Y} = \frac{\mathcal{T}}{Y}$ and $\bar{b} = 0$. So,

$$
\beta\widehat{b}_t^y = \widehat{b}_{t-1}^y - \bar{t}(\widehat{\tau}_t + \widehat{\omega}_t + \widehat{l}_t)
$$

with $\bar{t} \equiv \frac{\mathcal{T}}{Y}$ and $\widehat{b}_t^y \equiv \frac{\Delta B_t/P_t}{Y}$.

$$
\begin{aligned}
\widehat{\tau}_t + \widehat{\omega}_t + \widehat{l}_t &= \widehat{\tau}_t + (\sigma\widehat{y}_t + \varphi\widehat{l}_t + \frac{\tau}{1-\tau}\widehat{\tau}_t) + \widehat{l}_t \\
&= \frac{1}{1-\tau}\widehat{\tau}_t + \sigma\widehat{y}_t + (\varphi+1)\widehat{l}_t \\
&= \frac{1}{1-\tau}\widehat{\tau}_t + \sigma\widehat{y}_t + \frac{\varphi+1}{1-\alpha}\Big(\widehat{y}_t - a_t\Big) \\
&= \frac{1}{1-\tau}\widehat{\tau}_t + \Big(\sigma + \frac{\varphi+1}{1-\alpha}\Big)\widehat{y}_t - \frac{\varphi+1}{1-\alpha}a_t \\
&= \frac{1}{1-\tau}\widehat{\tau}_t + \Big(\sigma + \frac{\varphi+1}{1-\alpha}\Big)\Big(\widehat{\widetilde{y}}_t + \widehat{y}_t^e\Big) - \frac{\varphi+1}{1-\alpha}a_t
\end{aligned}
$$

$$\Rightarrow \widehat{b}_t^y = \beta^{-1}\widehat{b}_{t-1}^y - \bar{t}\beta^{-1}\Big[\frac{1}{1-\tau}\widehat{\tau}_t + \Big(\sigma + \frac{\varphi+1}{1-\alpha}\Big)\Big(\widehat{\widetilde{y}}_t + \widehat{y}_t^e\Big) - \frac{\varphi+1}{1-\alpha}a_t\Big]$$

Transversality condition:

$$\lim_{T\to\infty}\beta^T E_t\{Q_T B_T\} = 0 \iff$$
$$\lim_{T\to\infty} E_t\{\beta^{T+1}e^{\widehat{q}_T}\frac{B_T/P_T}{Y}P_T Y\} = 0 \approx \lim_{T\to\infty}\beta^{T+1}E_t\{\widehat{b}_T^y\} = 0$$

The optimization problem in this case writes

$$\min_{\pi_t,\widehat{\widetilde{y}}_t,\widehat{i}_t,\widehat{\tau}_t,\widehat{b}_t^y} E_0\sum_{t=0}^{\infty}\beta^t\Big[-\Lambda\widehat{\widetilde{y}}_t + \frac{1}{2}\Big(\pi_t^2 + \nu\widehat{\widetilde{y}}_t^2\Big)\Big]$$

$$\begin{aligned}
&\widehat{\widetilde{y}}_t = E_t\{\widehat{\widetilde{y}}_{t+1}\} - \frac{1}{\sigma}\Big(\widehat{i}_t - E_t\{\pi_{t+1}\} - \widehat{r}_t^e\Big)\\
&\pi_t = \beta E_t\{\pi_{t+1}\} + \lambda\Big(\sigma + \frac{\alpha+\varphi}{1-\alpha}\Big)\widehat{\widetilde{y}}_t + \lambda\frac{\tau}{1-\tau}\widehat{\tau}_t\\
&\widehat{b}_t^y = \beta^{-1}\widehat{b}_{t-1}^y - \bar{t}\beta^{-1}\Big[\frac{1}{1-\tau}\widehat{\tau}_t + \Big(\sigma + \frac{1+\varphi}{1-\alpha}\Big)(\widehat{\widetilde{y}}_t + \widehat{y}_t^e) - \frac{1+\varphi}{1-\alpha}a_t\Big]\\
&\lim_{T\to\infty}\beta^{T+1}E_t\{\widehat{b}_T^y\} = 0\\
&\Big(\widehat{i}_t + \rho \geq 0\Big)
\end{aligned}$$

with $\widehat{b}_t^y \equiv \frac{\Delta B_t/P_t}{Y}$. The optimality conditions write

$$\begin{aligned}
\pi_t :&\pi_t - (\sigma\beta)^{-1}\lambda_{t-1}^1 + \lambda_t^2 - \lambda_{t-1}^2 = 0\\
\widehat{\widetilde{y}}_t :&-\Lambda + \nu\widehat{\widetilde{y}}_t + \lambda_t^1 - \beta^{-1}\widehat{\lambda}_{t-1} - \lambda\Big(\sigma + \frac{\alpha+\varphi}{1-\alpha}\Big)\lambda_t^2 + \bar{t}\beta^{-1}\Big(\sigma + \frac{1+\varphi}{1-\alpha}\Big)\lambda_t^3 = 0\\
\widehat{i}_t :&\frac{1}{\sigma}\lambda_t^1 + \lambda_t^5 = 0\\
\widehat{\tau}_t :&-\lambda\frac{\tau}{1-\tau}\lambda_t^2 + \frac{\bar{t}\beta^{-1}}{1-\tau}\lambda_t^3 = 0\\
\widehat{b}_t^y :&\lambda_t^3 - \lambda_{t+1}^3 = 0, \forall t \neq T,\quad \lambda_T^3 + \lambda_T^4 = 0\\
&\Big(\widehat{i}_t + \rho\Big)\lambda_t^5 = 0,\quad \lambda_t^5 \leq 0, \widehat{i}_t + \rho \geq 0
\end{aligned}$$

alongside all equality constraints. Figure 3.23 shows the impulse responses under optimal monetary-fiscal policy mix for three different (small) levels of real transfers-to-output ratios τ. For $\tau = 0$, the model is isomorphic to the basic New-Keynesian one with market power distortions corrected by an employment subsidy financed by lump-sum taxes on households. Thus, in this case the optimal policy response is a one-to-one variation of the policy rate with respect to the efficient real interest rate. Under this policy, the economy is completely insulated from the effects of the demand preference shock, and hence there is no variation in equilibrium in any variable apart from the policy rate.

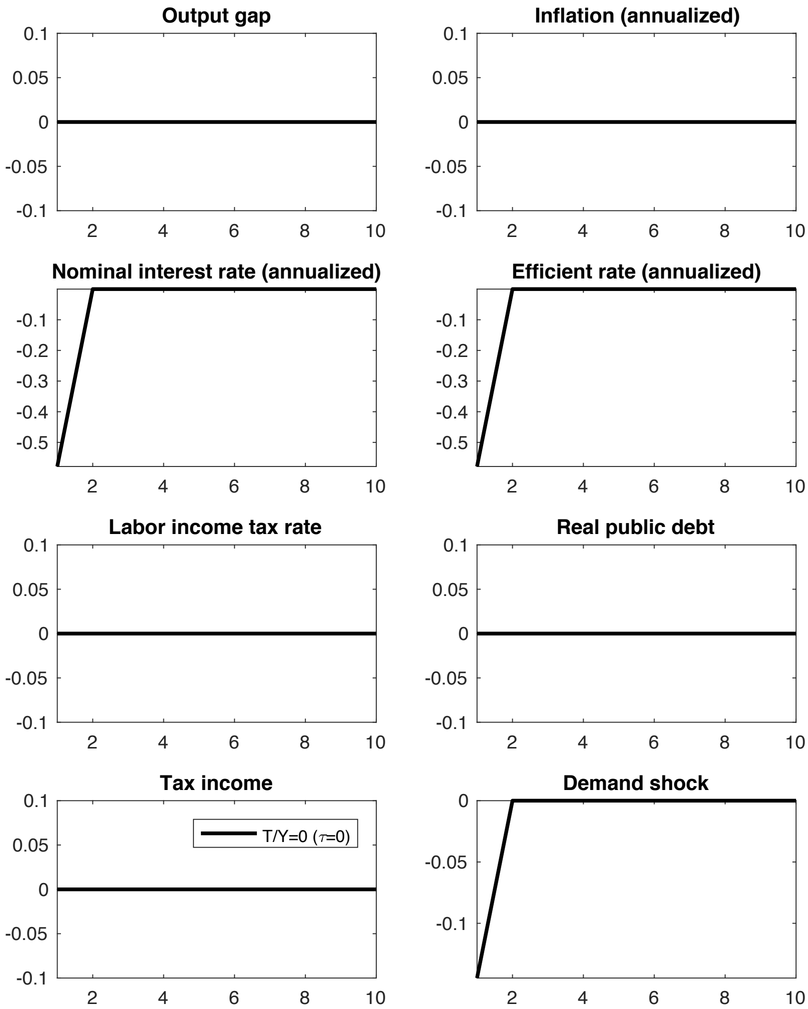


Figure 3.22: Dynamic responses to a transitory negative demand shock under optimal policy ($\bar{b} = 0$, zero steady-state transfers)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters

Strictly positive values of τ imply steady-state distortions as a result of the use of distortionary taxes to finance such transfers. It thus become optimal for the policymaker to engineer a persistent positive variation in the output gap in the aftermath of the shock so as to reduce these distortions. At the small level of real transfers and the associated low levels of (distortionary) tax rates, it turns out to be efficient to exclusively use the tax instrument to correct for such distortions. Specifically, the policy rate is varied one for one with the efficient interest rate so as to perfectly counteract the effects of the shock on demand, whereas tax rates are increased in response to the (unexpected) demand disturbance so as engineer an increase in tax income and hence an unexpected initial decline in real public debt. This initial strong decline in real public debt pushes inflation and output gap upwards in equilibrium. And, most importantly, it allows future tax rates to remain persistently below their steady-state levels, and hence output to persistently remain above its long-run level in the aftermath of the shock.

Note that in this case with a zero steady-state debt level, variations in inflation and interest rates do not have any first order effects on real public debt developments. Thus, the increase in the labor income tax alone is the one triggering the decline in real public debt. In particular, the decline in real public debt is not due to a unexpected inflation surprise by the policymaker.

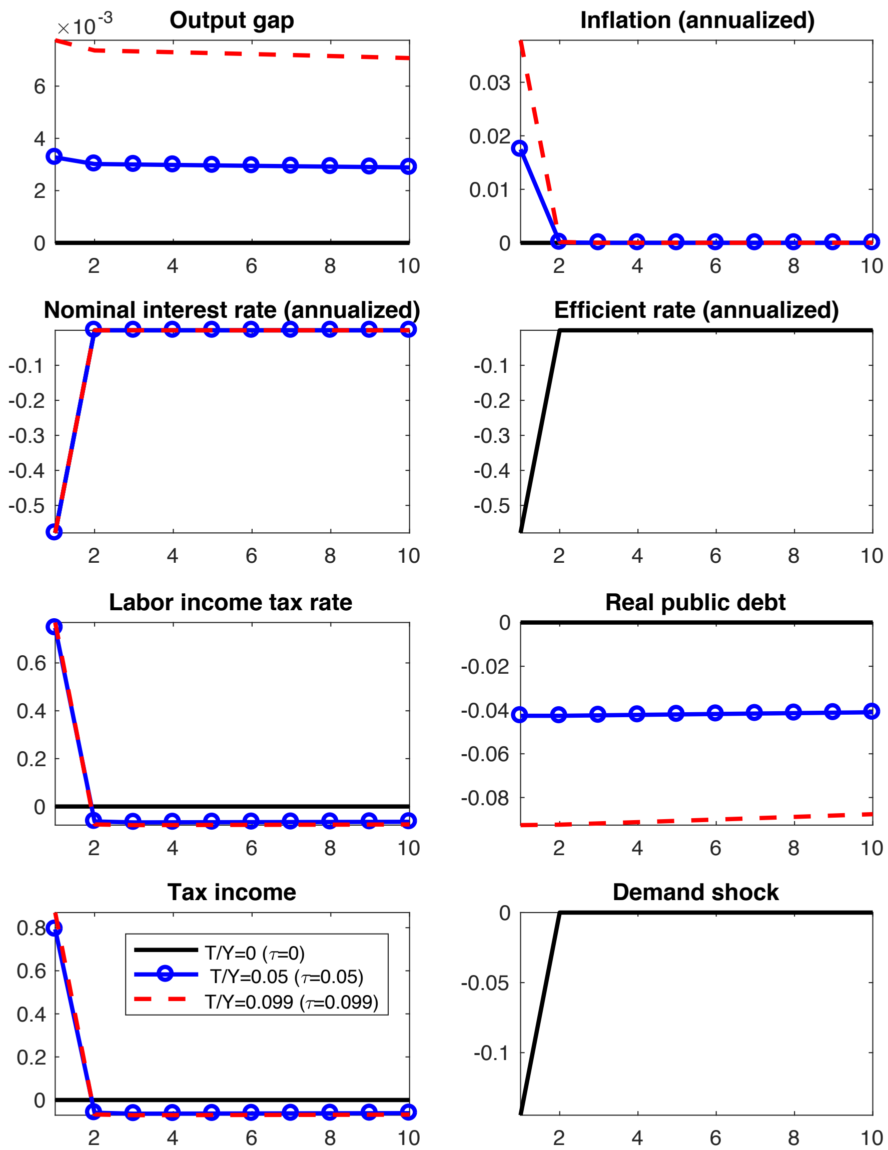


Figure 3.23: Dynamic responses to a transitory negative demand shock under optimal policy ($\bar{b} = 0$, positive steady-state transfers)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

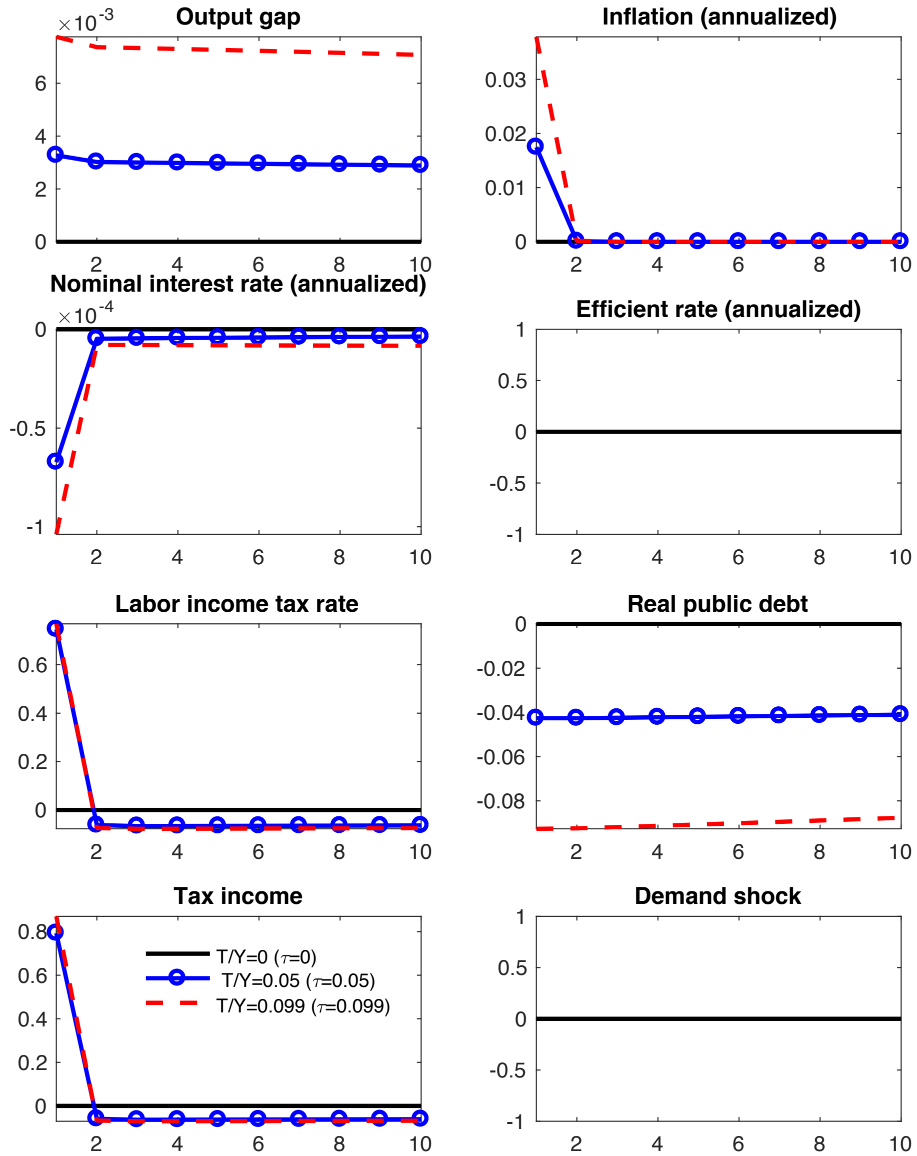


Figure 3.24: Optimal policy without shocks ($\bar{b} = 0$, positive steady-state transfers)
Note: Y-axis: % deviation from steady-state. X-axis: quarters

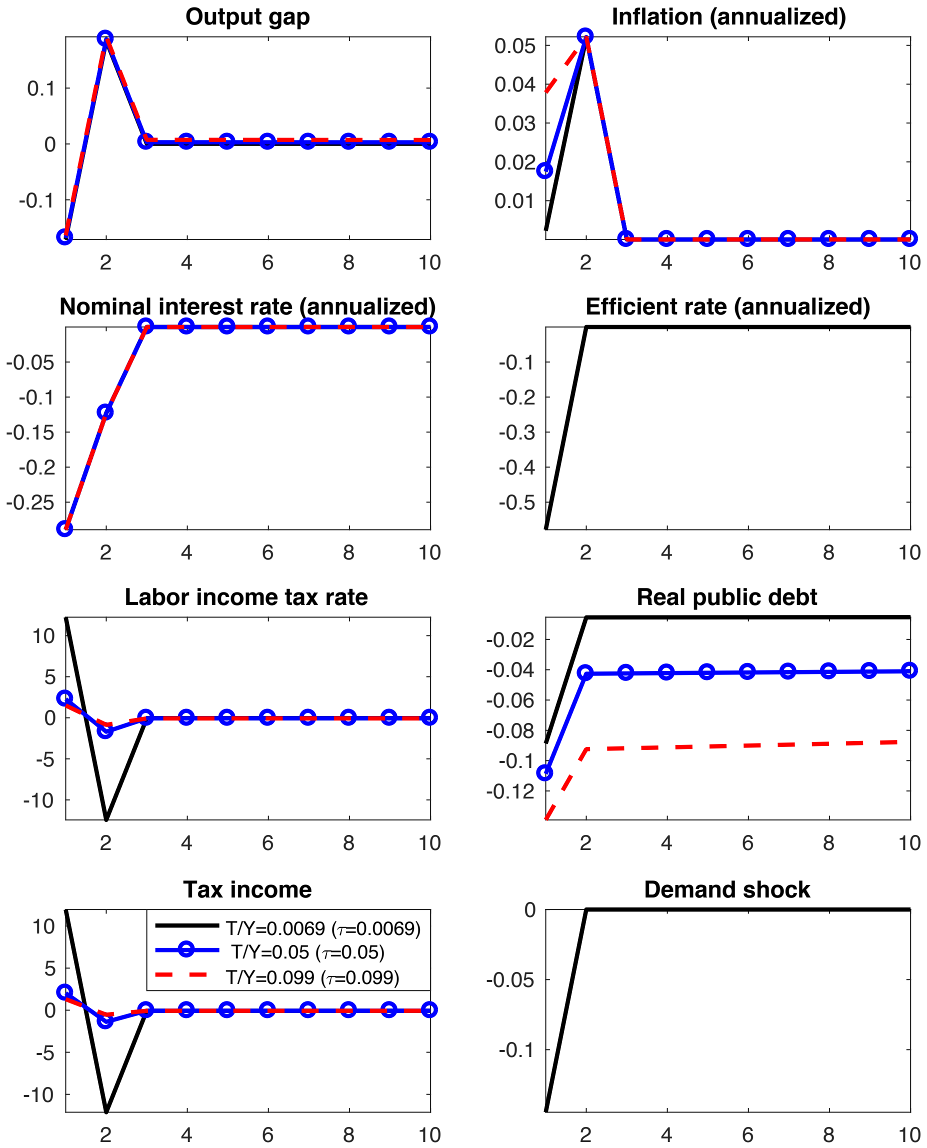


Figure 3.25: Dynamic responses to a transitory negative demand shock under optimal policy (small steady-state distorsions, $\bar{b} = 0$, ZLB)
Note: Y-axis: $\%$ deviation from steady-state. X-axis: quarters