The London School of Economics and Political Science

Essays on Macroeconomic Fluctuations

Thomas Drechsel

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy, London, September 2018

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it). The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

A version of Chapter 2 of this thesis was published in the *Journal of International Economics*, Vol. 112, May 2018, pp. 200-2018.

A version of Chapter 4 was published in the *Review of Economics and Statistics*, Vol. 99(2), May 2017, pp. 343-256.

I declare that my thesis consists of approximately 40,000 words, excluding graphs, tables and appendices.

Statement of conjoint work

Chapter 2 is joint work with Silvana Tenreyro and I contributed 50% of this work.

Chapter 3 is joint work with Wouter Den Haan and I contributed 50% of this work.

Chapter 4 is joint work with Juan Antolin-Diaz and Ivan Petrella and I contributed one third of this work.

Abstract

This thesis investigates fluctuations in the macroeconomy, from both empirical and theoretical angles, and in the context of developed as well as emerging economies.

Chapter 1 focuses on the role of firm borrowing for macroeconomic fluctuations in the United States. It presents micro-level evidence which highlights that firms' access to debt is constrained by their current earnings. Such a constraint leads to predictions about the transmission of investment shocks that are different from a traditional collateral constraint. The chapter tests these predictions using both aggregate and firm-level data. Empirical dynamics in the times series and cross-sectional dimension strongly support the relevance of the earnings-based borrowing constraint.

Chapter 2 turns to an open economy context and tackles the question of how important movements in international commodity prices are for emerging economy boom and bust cycles. For the case of Argentina, the chapter quantifies this nexus and finds a sizeable influence of commodity price shocks for movements in output, consumption and investment.

Chapter 3 demonstrates that misspecification of macroeconomic models can have severe consequences when estimating those models on the data. It proposes a novel concept to alleviate this concern, so-called agnostic structural disturbances (ASDs). The idea behind ASDs is to enrich the empirical specification of models while relying on relatively loose assumptions about how this restricts the dynamics.

Chapter 4 is concerned with tracking economic activity in real time. It develops a dynamic factor model that allows for changes in both the long-run growth rate of output and the volatility of business cycles. It documents a significant decline in long-run output growth in the United States most of which occurred prior to the Great Recession. The proposed model is capable of detecting shifts in long-run growth in a timely and reliable manner.

Acknowledgements

I would like to thank my advisors at the LSE for guiding me through the process of writing this thesis. Silvana Tenreyro, my main PhD supervisor, has not only taught me a great deal about economics, but also about the subtleties of communicating research and about writing concisely. She is a truly impressive person and working with her has been inspiring in many ways. I have acquired a lot of the skills I use in this thesis from being taught by and working with Wouter Den Haan. His patience and attention to detail have improved all parts of this thesis at all stages. Meetings with him were always extremely pleasant, both because we had many deep discussions about macroeconomics and because of his sense of humor. Ricardo Reis, who joined the LSE when I was roughly midway through the PhD, has contributed to this thesis with many incredibly sharp comments. A twenty minute meeting with him often felt like fast-reading an entire graduate textbook on macroeconomics. I very much hope that I can benefit from further interactions with the three of them in the future.

I owe a debt of gratitude to my collaborators Juan Antolin-Diaz and Ivan Petrella, with whom I have co-authored Chapter 4 of this thesis and who continuously provided extremely useful input on the other chapters. I have learned a lot about economics from both of them. Without the countless discussions I had with Juan at various pubs and restaurants around London, my PhD experience would not have been what it was, at an intellectual and a personal level. I will miss those days.

This thesis has also benefited from the generous help of various faculty members and students at the LSE, and from comments given by many participants of conferences and seminars in which I presented parts of this thesis. I especially thank Philippe Aghion, Andrea Alati, Neele Balke, Miguel Bandeira, Charlie Bean, Matteo Benetton, Kai Barron, Adrien Bussy, Francesco Caselli, Laura Castillo, Andy Ek, Simon Gilchrist, James Graham, Ben Hartung, Kilian Huber, Jay Hyun, Ethan Ilzetzki, Rustam Jamilov, Felix Koenig, Sevim Koesem, Per Krusell, Andrea Lanteri, Ralph Luetticke, Davide Melcangi, Jonathan Pinder Lukasz Rachel, Soren Ravn, Claudia Robles-Garcia, Philipp Schnabl, Roman Sustek, Horng Wong and Shengxing Zhang for very valuable suggestions. I have generally experienced the atmosphere among the LSE PhD students to be very stimulating, and some of my class mates have become friends for life. Thanks, gang!

I am infinitely happy that despite the huge time investment that was necessary to write this thesis, I managed to maintain my dearest friendships outside of the PhD. I am truly grateful to Andy and Holger for distracting me from research when I needed it, and to Simon, Daniel and Philip for visiting me in London on so many intense weekends. Finally, I thank my parents, my sister and my nephew, who was born around the time when Chapter 2 of this thesis was published. My work is dedicated to them.

Contents

List of Tables

List of Figures

Chapter 1

Earnings-Based Borrowing Constraints and Macroeconomic Fluctuations

1.1 Introduction

Firm credit displays large cyclical swings which correlate with fluctuations in output, employment and investment. Research on the drivers of this comovement has focused on how constraints to credit evolve over the business cycle and how this feeds back to economic activity.[1] This paper studies the macroeconomic consequences of an earnings-based constraint on firm borrowing. The focus on such a constraint is motivated by direct evidence on the importance of firms' current earnings flows for their access to debt. Micro data covering more than 50,000 loans to 15,000 US companies reveals the pervasive use of earnings-based loan covenants that make it difficult for firms to borrow when their current earnings are low.[2] I show that incorporating the earnings-based constraint in business cycle analysis is crucial for correctly capturing aggregate and firm-level credit dynamics and for characterizing the transmission of fiscal and monetary policy in the macroeconomy.

Earnings-based borrowing constraints imply dynamics of firm debt that are different from the ones generated by asset-based collateral constraints, which have become a cornerstone of many business cycle models that incorporate credit.[3] To demonstrate this, I build a theoretical model in which firm debt can be restricted either by a multiple of the firm's current earnings or by a fraction of the expected value of its

[1]Gertler and Gilchrist (2018) survey how research on the role of credit in macroeconomics has evolved since the 2008-09 global financial crisis, an event that marked a strong revival of this research agenda.

[2]My motivating evidence builds on existing empirical studies in corporate finance, in particular Lian and Ma (2018) who propose earnings-based constraints as a key determinant for firms' access to debt.

[3]The catalyst for research on collateral constraints was the seminal work by Kiyotaki and Moore (1997). Examples include Jermann and Quadrini (2012) who study firm borrowing in a macroeconomic context.

capital.[4] Under the alternative constraints, firm debt responds with opposite sign to structural shocks that move current earnings and the expected value of collateral in different directions. This is the case for a positive investment shock, which improves the ability of firms to turn resources into productive investment. Such a shock causes stronger economic activity and larger earnings, while it reduces the relative value of capital. Larger earnings allow for more debt under the earnings-based constraint, whereas, in contrast, the lower market value of capital reduces debt access with the collateral constraint.[5] In other words, to the extent that fluctuations are driven by shocks to firm investment, earnings-based constraints imply a positive comovement of firm debt with the business cycle, whereas collateral constraints imply a negative one.

The corresponding dynamics of debt in the data are consistent with the earnings-based constraint, and not in line with the predictions implied by a collateral constraint. To verify the model predictions, I study the dynamics of debt, earnings and capital following investment shocks in both macroeconomic and firm-level US data. In addition to exploiting the sharp differential predictions under the alternative constraints, there are two key advantages of my focus on investment shocks. First, previous studies have highlighted that shocks to the investment margin of the economy are an important quantitative feature of business cycles (see for example Justiniano, Primiceri, and Tambalotti, 2010, 2011).[6] Second, investment shocks can be identified in the data, based on a well-defined empirical counterpart, the inverse of the relative price of investment goods. I exploit identification strategies using this observable to verify the differential model predictions stemming from earnings-based versus collateral constraints.

Debt dynamics in aggregate data support the relevance of earnings-based borrowing constraints for the economy as a whole. This finding is based on a structural vector autoregression (SVAR) for US time series. I identify investment shocks using two alternative identification schemes, imposing long-run restrictions (following Fisher, 2006), as well as medium-run restrictions (in the spirit of Barsky and Sims, 2012). In both cases, the shock is identified based on its low frequency impact on the inverse relative price of equipment investment. I find that aggregate business sector debt increases in response to a positive investment shock, supporting the economy-wide importance of earnings-based constraints, and inconsistent with the dynamics implied by collateral constraints. In line with the model mechanism, business earnings rise and the value of the capital stock falls.

[4] I also discuss microfoundations for the presence of "asset-based" and "cash flow-based" lending.

[5] As an illustration, think about an airline and imagine a shock that makes the production of airplanes more efficient and lowers their relative price. The implication of this shock for borrowing differs sharply depending on the constraint. If airplanes serve as collateral, their falling relative value tightens the borrowing constraint. By contrast, the earnings-based constraint is relaxed as cheaper airplanes increase the airline's profitability.

[6] My use of the term investment shock at this stage encompasses different variations of this shock, including investment-specific technology shocks, marginal efficiency of investment shocks, as well as shocks to investment adjustment costs. I provide more detail on the differences between these concepts throughout the paper.

Heterogeneous debt dynamics in firm-level data provide further support for the model mechanism. I classify firms into those that face earnings-based covenants and those that borrow against collateral, and study their heterogeneous responses to investment shocks. Using a panel-version of the local projection method of Jordà (2005), I regress individual firm borrowing on the macro investment shock estimated from the SVAR, interacted with dummy variables that indicate earnings or collateral borrowers. To address endogenous selection into borrower types I control for rich firm characteristics and use different fixed-effect specifications. The results show that earnings-based borrowers significantly and persistently increase borrowing in response to a positive investment shock. The response of collateral borrowers is either negative or flat depending on the specification.[7]

Finally, earnings-based borrowing constraints also alter quantitative conclusions about US business cycles, in particular the transmission of policy shocks. I extend my model to incorporate features of a New Keynesian structural macroeconomic model. The extended model features a number of additional shocks and frictions, such as price and wage rigidities, alongside a constraint which limits debt by a combination of an earnings-based and a collateral component. I estimate the weight between the components as well as other structural parameters of the model on US postwar data. The data assigns a posterior weight of 0.9 on the earnings-based borrowing component, 0.1 on the collateral component. Counterfactual estimations indicate that the presence of earnings-based constraints dampens the output response of fiscal shocks, whereas monetary shocks have much stronger effects on inflation and somewhat stronger but less persistent effects on output. The intuition for the former result is that fiscal shocks crowd out investment to a larger extent when there is no additional benefit from building up collateral. The latter result is driven by a low degree of estimated price rigidity that emerges in the presence of the earnings-based constraint. The estimated model also implies that investment shocks more than half of the variation in US output, which lends further support to my focus on this shock for verifying the relevance of earnings-based borrowing constraints in macro and micro data.

Relation to the literature. First and foremost, this paper contributes to the vast literature on the role of financial frictions in macroeconomics, going back to the seminal work of Bernanke and Gertler (1989), Shleifer and Vishny (1992), and

[7] In a formal test, the null hypothesis of equal responses across borrower types is rejected. In an alternative setting I also show that firm-level responses of debt to a fall in the relative price of investment goods, using investment shocks as an instrumental variable (IV), are consistent with the proposed mechanism.

Kiyotaki and Moore (1997).[8,9] In a retrospective on business cycle models, Kehoe, Midrigan, and Pastorino (2018) highlight the importance of disciplining macro models with direct micro evidence. In this spirit, my paper uses micro evidence on firm borrowing to capture firm debt dynamics more accurately in the context of studying macroeconomic fluctuations.

Second, the motivating evidence I provide builds on existing insights, highlighted by the empirical corporate finance literature, on the pervasive use of loan covenants. Important contributions are Chava and Roberts (2008) and Sufi (2009).[10] The focus of my paper is closely related to that of Lian and Ma (2018) who investigate the relevance of cash flow-based relative to asset-based firm borrowing. Based on a comprehensive empirical analysis, their paper proposes that the key constraint to firm debt are cash flows measured by earnings. These authors mainly focus on causally identifying the extent to which increases in earnings relax borrowing constraints at the micro level.[11] My paper embeds the relevance of borrowing based on current earnings into a macroeconomic model, verifies the predicted dynamics empirically at both the micro and the macro level, and demonstrates quantitative consequences for business cycles.

Third, my model predictions and their empirical verification relate to the literature on investment shocks, which includes theoretical work by Greenwood, Hercowitz, and Huffman (1988) and Greenwood, Hercowitz, and Krusell (2000), and papers that identify investment shocks in SVARs building on the key contribution by Fisher (2006). Justiniano, Primiceri, and Tambalotti (2010, 2011) investigate the role of investment shocks in US business cycles and find them to be a key force behind output fluctuations. I contribute to this literature by analyzing borrowing dynamics that arise from investment shocks.[12,13]

[8]Research on borrowing constraints includes for example Lorenzoni (2008), Geanakoplos (2010), Gertler and Karadi (2011), Kiyotaki and Moore (2012), Jermann and Quadrini (2012), Liu, Wang, and Zha (2013), Buera and Moll (2015), Azariadis, Kaas, and Wen (2016), Del Negro, Eggertsson, Ferrero, and Kiyotaki (2016) and Cao and Nie (2017). Kocherlakota (2000) and Cordoba and Ripoll (2004) investigate the quantitative importance of collateral constraints. While collateral constraints are typically based on limited contract enforcement, another important class of financial frictions is based costly state verification problems in the spirit of Townsend (1979), see e.g. Bernanke and Gertler (1995), Carlstrom and Fuerst (1997) and Bernanke, Gertler, and Gilchrist (1999). Quadrini (2011) provides a survey on financial frictions in macroeconomics.

[9]Recent papers that focus on corporate debt dynamics over the business cycle but do not highlight earnings-based constraints include Crouzet (2017), Xiao (2018) and Grjebine, Szczerbowicz, and Tripier (2018).

[10]Other papers that focus on covenants include Dichev and Skinner (2002), Roberts and Sufi (2009a, 2009b), Nini, Smith, and Sufi (2012), Murfin (2012), Bradley and Roberts (2015) and Falato and Liang (2017). Chodorow-Reich and Falato (2017) study empirically how bank health transmitted to the economy via covenants during the 2008-09 financial crisis. For a theoretical treatment see Garleanu and Zwiebel (2009).

[11]Their paper also contains an exploration of cash-flow based lending in a Kiyotaki-Moore economy. An earlier paper that aims to identify the determinants of borrowing constraints at the micro level, but does not focus on earnings constraints, is Chaney, Sraer, and Thesmar (2012).

[12]See also Schmitt-Grohe and Uribe (2012) for a business cycle model with investment shocks. Other papers in the SVAR literature include Barsky and Sims (2012) and Francis, Owyang, Roush, and DiCecio (2014).

[13]My econometric approach to studying firm-level responses to investment shocks using local projections relates to work by Jordà, Schularick, and Taylor (2017) and Cloyne, Ferreira, Froemel, and Surico (2018).

Fourth, there are a few existing papers, within and outside the business cycle literature, in which flow variables rather than assets restrict borrowing (for example Kiyotaki, 1998).[14] My contribution relative to these papers lies in explicitly comparing theoretical and empirical differences between income flow-related and collateral constraints on firms.[15] I provide a detailed exploration of how different stock and flow borrowing constraints relate and demonstrate that the definition of earnings as opposed to other financial flows is key for characterizing empirically plausible debt dynamics with the earnings-based constraint.

Fifth, in many countries mortgage contracts also contain income-related constraints, often directly imposed by the regulator. Greenwald (2017) formulates a payment-to-income limit in addition to a collateral (loan-to-value) constraint for mortgage borrowing and studies the transmission of macroeconomic shocks through the mortgage market.[16] My paper focuses on corporate debt rather than household mortgages, where the relevance of earnings-based constraints for business cycles is still understudied.

Finally, in a recent paper Adler (2018) studies the aggregate impact of financial covenants on investment in a business cycle model. Relative to his work, I focus less explicitly on the aspect of covenant breaches, but interpret the prevalence of earnings covenants as evidence of an earnings-based credit constraint. Moreover, I derive and verify predictions for debt dynamics at the macro and micro level and investigate consequences of the constraint for the transmission of monetary and fiscal policy.

Structure of the paper. Section 1.2 presents microeconomic evidence that motivates the focus on earnings-based borrowing constraints for firms. Section 1.3 introduces a business cycle model that features an earnings-based constraint and discusses the resulting debt dynamics in comparison to a collateral constraint. Section 1.4 verifies the differential theoretical predictions for investment shocks using both SVAR analysis on aggregate data and panel projections on firm-level data. Section 1.5 turns to quantitative questions by estimating a New Keynesian model with earnings-based borrowing. Section 1.6 concludes.

[14] See also Jappelli and Pagano (1989) in the context of the permanent income hypothesis and Arellano and Mendoza (2002), Mendoza (2006), Bianchi (2011) and Korinek (2011) in the context of sovereign debt. Brooks and Dovis (2018) examine the sensitivity of credit constraints to profit opportunities in a trade framework. Li (2016) studies how the lack of pledgeability of both assets and earnings reduces aggregate productivity in Japan.

[15] Constraints based on income flows often provide an ad-hoc way to restrict borrowing, for example if the model does not feature capital. Schmitt-Grohe and Uribe (2016) provide some discussion on stock versus flow constraints on sovereign debt. Diamond, Hu, and Rajan (2017) lay out a theory of firm financing in which control rights both over asset sales and over cash flows have varying importance over time.

[16] A related study is Corbae and Quintin (2015). Earlier work that studies household mortgages in business cycle models typically focuses on collateral, see for example Iacoviello (2005) and Iacoviello and Neri (2010).

1.2 Motivating evidence on earnings-based borrowing

This section presents stylized evidence on corporate borrowing in the US economy. Using information from more than 50,000 loan deals issued to 15,000 firms, I document that earnings are a key indicator that determines the extent to which firms have access to loans.

Data source. I use the *ThomsonReuters LPC Dealscan* data base. For the United States, this data covers around 75% of the total commercial loan market in terms of volumes.[17] The unit of observation is a loan *deal*, which consists of loan *facilities*. Deal and facility can be the same unit, e.g. for a standard bank loan, or a deal can consist of a syndicated credit arrangement in which several lenders provide facilities of different types and conditions. The data contain rich information, including the identity of borrower and lender, the amount, maturity, and interest rate. I consider USD denominated loan originations since 1994 for US nonfinancial corporations. In Section 1.4.3, I merge the Dealscan data to the *Compustat Quarterly* data, which covers accounting information of listed US companies.[18]

The pervasive use of loan covenants. Loan covenants, sometimes referred to as nonprice terms, are legal provisions which the borrower is obliged to fulfill during the lifetime of a loan. They are usually linked to specific measurable indicators, for which a numerical maximum or minimum value is specified. A covenant states for example that "the borrower's earnings-to-debt ratio must be above 4". Covenant breaches lead to technical default, which gives lenders discretion in taking contingent actions: calling back the loan, imposing a penalty payment, increasing the interest rate or changing other conditions in the contract. Breaches have been shown to occur frequently with large economic effects. Chodorow-Reich and Falato (2017) show for example that one third of nonfinancial firms breached their covenants during the 2008-09 financial crisis. Importantly, Roberts and Sufi (2009a) find that net debt issuing activity experiences a large and persistent drop immediately after a covenant violation.[19] These findings indicate that debt access is significantly reduced when the variable specified in the covenant moves above (below) its maximum (minimum) value.

The importance of earnings. Table 1.1 lists the most popular covenant types, sorted by their frequency of use. The frequency is calculated for loans that feature at least one covenant and the table includes covenants which appear in more than 10% of these loans. Note that a given contract in the sample can have up to eight covenants. The table also contains the median, 25th and 75th percentile as well as

[17] See Chava and Roberts (2008). The data does not include a big share of marketable debt instruments such as corporate bonds, a limitation which I will discuss later in this section.

[18] Appendix 1.7.1 contains further information on the data set as well as summary statistics.

[19] Chava and Roberts (2008) find strong effects of breaches on investment and Falato and Liang (2017) strong effects on employment.

the value-weighted mean of the covenant value, that is, the numerical maximum or minimum value that restricts a given indicator.

Table 1.1: LOAN COVENANT TYPES, VALUES AND FREQUENCY OF USE

	Covenant type	**p25**	**Median**	**p75**	**Mean**	**Frequency**
1	Max. Debt to EBITDA	3.00	3.75	5.00	4.60	60.5%
2	Min. Interest Coverage (EBITDA / Interest)	2.00	2.50	3.00	2.56	46.7%
3	Min. Fixed Charge Coverage (EBITDA / Charges)	1.10	1.25	1.50	1.42	22.1%
4	Max. Leverage ratio	0.55	0.60	0.65	0.64	21.3%
5	Max. Capex	6M	20M	50M	194M	15.1%
6	Net Worth	45M	126M	350M	3.2B	11.5%

Note: The table list the most pervasive covenant types, sorted by their frequency of use in the Dealscan loan data. Covenant types with a frequency above 10% are included in the table. As there can be more than one covenant per loan, the frequency adds up to more than 100%. EBITDA abbreviates earnings before interest, taxes, depreciation and amortization. As indicated in brackets, a minimum interest coverage covenant typically links to the ratio of EBITDA to interest expenses and a minimum fixed charge coverage covenant to the ratio of EBITDA to fixed loan charges. The sample consists of loan deals with at least one loan covenant, issued between 1994 and 2015 by US nonfinancial corporations. The mean and frequency are weighted with real loan size. 'M' and 'B' refer to million and billion of 2009 real USD, respectively.

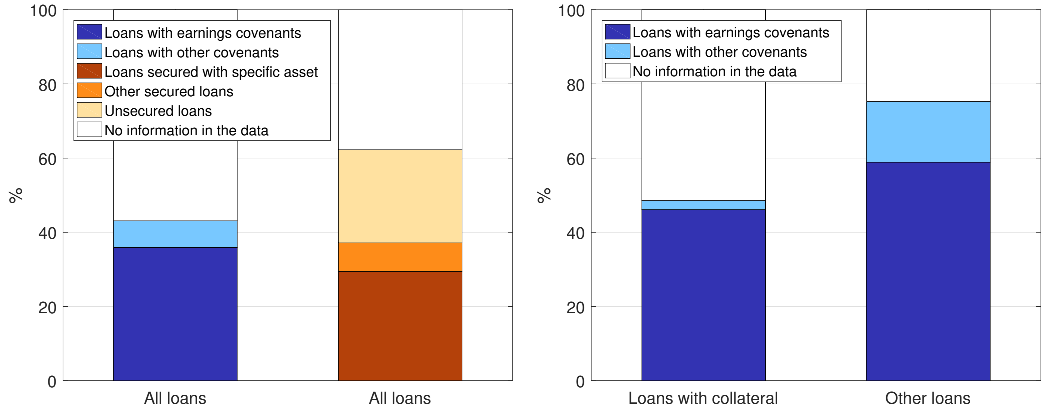
The table shows that the three most frequently used covenants are all related to earnings. The specific earnings measure is EBITDA, which measures earnings before interest, taxes, depreciation and amortization. EBITDA is a widely used indicator of a firm's economic performance. It captures firm profits that come directly from its regular operations and is not confounded by accounting treatment of taxes and depreciation. It is also readily available for scrutiny by lenders as part of standard financial reporting. The most frequently occurring covenant implies that the lender requires the level of debt not to exceed this measure by a multiple of 4.6 on average at any given point in time. At this stage, I interpret the prevalence of earnings-based covenants as suggestive evidence that the flow of current earnings constitutes an important constraint on companies' access to debt. The subsequent sections of this paper will be dedicated to studying whether credit dynamics support this interpretation, and whether it affects conclusions about aggregate fluctuations.

Further channels through which earnings affect debt access. Loan covenants are a direct manifestation of current earnings potentially constraining access to debt, as they are explicitly written into contracts. There is also evidence of *implicit* debt constraints related to earnings. For example, lenders may base their decisions on credit ratings, which are typically constructed with a strong emphasis on EBITDA. Furthermore, scrutiny of earnings by lenders could come in the form of internal credit risk models that use earnings as an input, or be based on reference levels in earnings ratios that lenders are accustomed to consider without explicitly using covenants.[20]

[20] According to Standard & Poor's Global Ratings (2013a, 2013b), the financial risk profile of corporations is assessed based on *core ratios*, which are the funds from operations (FFO)-to-debt and the debt-to-EBITDA ratio, as well as *supplemental ratios*, which relate to other operating cash

Figure 1.1: IMPORTANCE OF EARNINGS-BASED AND ASSET-BASED DEBT IN COMPARISON

(a) Frequencies of covenants and secured loans (b) Covenants within (un)collateralized loans



Note: Panel (a) displays the value-weighted shares of loan deals that contain covenants (left bar) and are secured/unsecured (right bar). In the left bar, the dark blue area represents the share with at least one earnings-based covenant. The light blue area covers loans with covenants unrelated to earnings. In the right bar, the different orange shades capture loans secured with specific assets (dark), other secured loans (medium) and unsecured loans (light). In both bars, loans without the relevant information are represented by the white area. Panel (b) repeats the left column of Panel (a), but breaks down the sample into loans secured with specific assets and other loans (with any information on secured/unsecured). The sample used for both panels consists of loan deals issued between 1994 and 2015 by US nonfinancial corporations.

Earnings-based vs. asset-based lending. Figure 1.1 analyzes the frequency of loan covenants and of collateral, that is, debt that is secured with specific assets.[21] Panel (a) plots different value-weighted shares in the total number of loans. The left bar presents the share of loans with at least one earnings-related covenant (dark blue area) and with only other covenant types (light blue area). For the remaining share, the information on covenants is not available (white area). The right bar presents the share of loans that are secured with specific assets, other secured loans, unsecured loans, and loans without information on whether and how they are secured (dark orange, medium orange, light orange, and white areas, respectively).[22] The left bar indicates that earnings-based covenants, which dominate within covenants overall, feature in around 35% of loans. This number is a lower bound, as the remainder of loans does not have any information, which does not necessarily mean that covenants are absent. The share of earnings-based covenants is higher than the share of debt secured by specific assets, shown in the right bar.[23] Note that other secured debt

flow measures. Together with the business risk profile (country risk, industry risk, competition) this determines the credit rating of a company.

[21] Since the information on secured/unsecured is at the facility-level, while the information on covenants is at the deal-level, I aggregate to the deal level, summing over the relevant facilities within deals.

[22] According to Lian and Ma (2018), loans secured with "all assets" in Dealscan should be classified as cash-flow based loans, as the value of this form of collateral in the case of bankruptcy is calculated based on the cash flows from continuing operations. Therefore, I define loans backed by specific assets as secured loans but exclude those that are backed by "all assets". I assign the latter to the separate category called "Other secured loans". I thank Yueran Ma and Chen Lian for a helpful discussion on these differences.

[23] Table 1.7 in the Appendix lists the frequencies of different asset classes that are used as collateral.

is composed of loans that are secured by the entire balance sheet of the borrower. Finally, it is noteworthy that a sizable chunk of loans are unsecured.

Panel (b) breaks down the frequency of covenants conditional on the loan being in two different groups. The first one is loans that are secured by specific assets while the second one is other loans, excluding loans without information on secured/unsecured. This provides evidence on the extent to which the use of loan covenants and collateral is related across loans. The panel shows that covenants overall are more likely to appear in a loan contract when specific collateral is not present. However, the loans backed by specific assets still have a reasonably high share of covenants. Taken together, the loan information suggests that earnings-based borrowing is pervasive, exceeding the prevalence of asset-based debt, and that earnings-based covenants are used both in addition to and instead of collateral.[24]

Existing evidence in the literature. Lian and Ma (2018) supplement the Dealscan data with a variety of data sources and present detailed evidence that US nonfinancial firms primarily borrow based on cash flows as measured by earnings. The magnitudes that Lian and Ma (2018) report paint a picture that is perhaps even stronger in favor of earnings-based borrowing, likely due to the fact that my analysis does not cover most marketable debt securities such as corporate bonds.[25] Other studies also resort to data sources different from Dealscan to study related questions. Azariadis, Kaas, and Wen (2016) use Compustat data to highlight the quantitative importance of firm borrowing without collateral.[26] While the focus on earnings-based constraints for firms is relatively novel, the fact that recent research using additional or other data sources reaches similar conclusions on the pervasiveness of such constraints lends further support to studying this microeconomic phenomenon in a business cycle context.

1.3 A business cycle model with earnings-based debt

This section proposes an earnings-based constraint on firm borrowing to formalize the microeconomic evidence. I set up a prototype business cycle model in which the firm issues one of two debt types, which are constrained by current earnings and the value of collateral, respectively. This allows me to study the dynamics arising from the earnings-based constraint in comparison with a traditional asset-

[24] Figure 1.15 in the appendix repeats the analysis with equal-weighted rather than value-weighted shares.

[25] This means that despite the comprehensive coverage of Dealscan within the universe of loans, a sizable chunk of aggregate corporate sector liabilities are not captured. To get a rough idea, I calculated using Flow of Funds data for 2016 that outstanding loans in the nonfinancial business sector amount to around 7.6 tn USD, while 5.8 tn USD of liabilities are in debt securities. The Dealscan data contains mostly syndicated loan deals and many of the facilities within a deal are credit lines (see also Appendix 1.7.1). Ivashina and Scharfstein (2010) study the syndication aspect of corporate borrowing in more detail. For work with a more explicit focus credit lines see for example Sufi (2009) and Acharya, Almeida, and Campello (2013).

[26] Azariadis, Kaas, and Wen (2016) use a specific Compustat item which captures secured debt and calculate unsecured debt as a residual, subtracting it from total debt liabilities.

based constraint. To derive sharp differential predictions, the characterization of the model dynamics focuses on a structural shock that moves earnings and the value of collateral in opposite directions: the investment shock. Section 1.5 extends the model to a quantitative framework and also highlights how the earnings-based constraint differentially affects the transmission of other shocks, including monetary and fiscal shocks.

1.3.1 Model environment

Time is discrete, denoted by t, and continues infinitely. The frequency is quarterly. The economy is populated by a representative firm and a representative household. There is a government which runs a balanced budget.

Firm problem

The firm produces a final consumption good using capital, which it owns and accumulates, and labor, which it hires on a competitive labor market taking the wage rate w_t as given. The consumption good is produced with a Cobb-Douglas production function

$$y_t = z_t k_{t-1}^{\alpha} n_t^{1-\alpha}, \tag{1.1}$$

and its price is normalized to 1. $\alpha \in (0,1)$ is the capital share in production. Total factor productivity (TFP), z_t, is subject to stochastic shocks, to be specified below. The firm's period earnings flow, or operational profits, is denoted as π_t and defined as

$$\pi_t \equiv y_t - w_t n_t. \tag{1.2}$$

This definition of earnings corresponds to *EBITDA*, that is, sales net of overhead and labor costs, but without subtracting investment, interest payments or taxes. Hence, the model definition in (1.2) is consistent with the indicator that features in the most pervasive covenant according to the evidence provided in Section 1.2. π_t is the measure that will enter the firm's earnings-based borrowing constraint.

Capital k_{t-1} is predetermined at the beginning of the period and its law of motion is

$$k_t = (1-\delta)k_{t-1} + v_t \left[1 - \Phi_t\left(\frac{i_t}{i_{t-1}}\right)\right] i_t, \tag{1.3}$$

where δ is the depreciation rate and v_t is a stochastic disturbance, following a process specified further below. In the environment presented here, where the production of consumption, investment and capital goods is not decentralized into different sectors, v_t captures both the level of the economy's investment specific technology (IST) as well as its marginal efficiency of investment (MEI). I refer to shocks to the process of v_t simply as "investment shocks".[27] The term $\Phi_t\left(\frac{i_t}{i_{t-1}}\right)$ introduces investment

[27] IST captures the efficiency at which consumption is turned into investment, while MEI represents the efficiency at which investment is turned into installed capital. Both types of disturbances have been studied extensively in business cycle research, e.g. by Greenwood, Hercowitz, and Krusell (2000) and Justiniano, Primiceri, and Tambalotti (2011). The key difference is that IST corresponds

adjustment costs. Following Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) I assume that $\Phi_t(1) = 0$, $\Phi_t'(1) = 0$, and $\Phi_t''(1) = \phi_t > 0$. The t subscript captures the possibility of stochastic shocks to adjustment costs. I refer to the composite term $v_t \left[1 - \Phi_t\left(\frac{i_t}{i_{t-1}}\right)\right]$ as the *investment margin*. The results I characterize below will show that for the purpose of disentangling the two alternative borrowing constraints, different types of shocks to this margin work in similar ways.

Both the presence of investment adjustment costs as well as v_t will lead to variation in the market value of capital. In the case of adjustment costs, this arises from the standard result that adjustment costs move the value of capital inside the firm relative to its replacement value, that is, they affect the ratio known as "Tobin's Q" (see for example Hayashi, 1982).[28] In the case of v_t, it is important to note that even in the absence of any adjustment costs, v_t will be inversely related to the relative price of k_t in consumption units. To see this, consider the flow of funds constraint of the firm, in units of the consumption good, which reads

$$\Psi(d_t) + i_t + b_{\pi,t-1} + b_{k,t-1} = y_t - w_t n_t + \frac{b_{\pi,t}}{R_{\pi,t}} + \frac{b_{k,t}}{R_{k,t}}. \tag{1.4}$$

$\Psi(d_t)$ denotes the dividend (equity payout) function, and the b terms capture debt financing, both of which will be explained in more detail below. Setting $\Phi_t(\cdot) = 0$ and substituting i_t from equation (1.3) into equation (1.4), it can be seen that the relative price of capital directly varies with the inverse of v_t, a key property of models that feature such disturbances entering the investment margin:

$$\Psi(d_t) + \frac{k_t}{v_t} + b_{\pi,t-1} + b_{k,t-1} = y_t - w_t n_t + \frac{(1-\delta)k_{t-1}}{v_t} + \frac{b_{\pi,t}}{R_{\pi,t}} + \frac{b_{k,t}}{R_{k,t}}. \tag{1.5}$$

This observation about the relative price of capital will play a key role in the dynamics of debt following investment shocks under different borrowing constraints.

The firm has access to two means of financing its operations, equity and debt. The variable d_t denotes equity payouts and the presence of the function $\Psi(d_t)$ captures costs related to equity payouts and issuance. Following Jermann and Quadrini (2012),

$$\Psi(d_t) = d_t + \psi(d_t - \bar{d})^2, \tag{1.6}$$

where $\bar{d}$ is the long run dividend payout target (the steady state level of d_t). Equation (1.6) captures in reduced form the fact that raising equity is costly and that there are motives for dividend smoothing.[29] Debt financing can be undertaken in the form of two alternative one-period risk-free bonds, denoted $b_{\pi,t}$ and $b_{k,t}$, where $b_{\pi,t-1}$ and

empirically to the inverse of the relative price of investment, while MEI does not have a readily available empirical counterpart. This will come into play when taking my model predictions to the data in Section 1.4.

[28] In his original paper, Hayashi (1982) introduces a similar formulation as in equation (1.3) and refers to the composite term $\left[1 - \Phi_t\left(\frac{i_t}{i_{t-1}}\right)\right] i_t$ as the "installation function". In his setting, there is no variation in IST and the price of investment goods in consumption units is exogenously given.

[29] Altinkilic and Hansen (2000) provide evidence of increasing marginal costs in equity underwriting. Lintner (1956) discusses dividend smoothing motives.

$b_{k,t-1}$ are predetermined at the beginning of period t. The effective gross interest rates faced by firms are $R_{\pi,t}$ and $R_{k,t}$, and are both subject to a tax advantage, captured by τ_π and τ_k, of the following form:

$$R_{j,t} = 1 + r_{j,t}(1 - \tau_j), \quad j \in \{\pi, k\} \tag{1.7}$$

where $r_{\pi,t}$ and $r_{k,t}$ are the market interest rates received by lenders. This creates a preference for debt over equity and makes the firm want to borrow up to its constraints. Since the household does not receive this tax rebate, there is a heterogeneity in the desire to borrow and save across sectors of the economy, the household wants to lend funds, and debt is in positive net supply in equilibrium. This type of tax exists in many countries and the related modeling assumption follows Hennessy and Whited (2005).[30]

Introduction of alternative borrowing constraints. Both types of debt are subject to borrowing constraints, which are formulated in consumption units and which I specify as

$$\frac{b_{\pi,t}}{1 + r_{\pi,t}} \leq \theta_\pi \pi_t \tag{1.8}$$

and

$$\frac{b_{k,t}}{1 + r_{k,t}} \leq \theta_k \mathbb{E}_t p_{kt+1}(1 - \delta)k_t. \tag{1.9}$$

The θ parameters capture the exogenous tightness of the constraints.[31] In the earnings-based borrowing constraint (1.8), debt is limited by a multiple $\theta_\pi > 1$ of current earnings, π_t.[32] I also allow a more general form of this constraint, in which $f(\pi_{t-3}, \pi_{t-2}, \pi_{t-1}, \pi_t, \mathbb{E}_t\pi_{t+1})$ enters on the right hand side, and $f(\cdot)$ is a linear polynomial. This captures the idea that loan covenant indicators in practice are typically calculated as 4-quarter trailing averages (see Chodorow-Reich and Falato, 2017). An alternative formulation of the earnings-based constraint would be one that captures the interest coverage ratio, that is, a constraint on $r_{j,t}b_{j,t}$. I focus exclusively on the debt-to-earnings formulation, as the corresponding covenant is the most frequently used in the loan data, ahead of the coverage ratio (see Table 1.1).[33]

[30]See also Riddick and Whited (2009). Strebulaev and Whited (2012) provide a comprehensive review of the dynamic corporate finance literature. In effect, the tax advantage makes the firm "less patient" than the market, which discounts at rate $\frac{1}{1+r_{j,t}}$, and thus induces the firm to borrow up to its constraint. This outcome could be generated in alternative ways, for example by making the firm an entrepreneur household with a lower discount factor.

[31]In Section 1.5, I allow these parameters to be time-varying and subject to stochastic "financial shocks" in the spirit of Jermann and Quadrini (2012).

[32]A constraint on $b_{j,t}$ rather than $\frac{b_{j,t}}{1+r_{j,t}}$ would capture a different timing of the interest payment and would not change the dynamics of the model in a meaningful way.

[33]Lian and Ma (2018) emphasize the presence of both debt-to-earnings ratios as well interest coverage ratios in covenants. Daniel Greenwald's discussion of their paper at the NBER Monetary Economics Meeting, available online, contains some further thoughts on the differences between the

In equation (1.9) debt issued by the firm in t is limited by a fraction $\theta_k < 1$ of the capital stock net of depreciation next period, which is valued at price $p_{k,t+1}$. In the borrowing constraint $p_{k,t+1}$ could reflect either the book or the market price of capital.[34] Formally,

$$p_{k,t} = \begin{cases} \frac{1}{v_t} & \text{if collateral is book value} \\ Q_t & \text{if collateral is market value} \end{cases} \tag{1.10}$$

where Q_t is the market price of capital, to be determined in equilibrium. In the presentation of the main results, I will focus on the market value formulation, but it is important to emphasize that in the presence of investment shocks the book price of capital is not 1 but $1/v_t$, as the debt contract is specified in consumption units. The equilibrium value of Q_t will also be inversely related to v_t but will be additionally affected by adjustment costs. If adjustment costs are set to zero, the market and book value of capital coincide at $1/v_t$.

Discussion of borrowing constraints. Borrowing constraints reflect that the ability of a borrower to issue debt is limited due to an underlying friction such as information or enforcement limitations. In the case of the collateral constraint, a large body of work shows how the market incompleteness implied by the constraint can be derived from such frictions. Typically, a collateral constraint emerges as the optimal solution in a setting in which borrowers have the ability to divert funds or withdraw their human capital from a project, but the withdrawal remains an off-equilibrium threat (see for example Hart and Moore, 1994).

In the case of the earnings-based borrowing constraint, one interpretation is that the firm is able to directly pledge its earnings stream rather than an asset in return for obtaining debt access. A second interpretation is that the borrower has the ability to divert funds, in which case the lender can seize and operate the firm herself. As the lender cannot perfectly predict the value of the firm when it is taken over, she estimates this contingent firm value as a multiple of current earnings.[35] A third interpretation is based on regulation. Regulatory requirements on lenders require a different risk treatment of loans that feature a low earnings-to-debt ratio. Exogenously imposed constraints that are not the outcome of a contracting problem could reflect such regulation.[36] In Appendix 1.7.3, I sketch out a specific formal environment that captures the second of these three potential interpretations. In that appendix, I also discuss the existing literature on the microfoundations of loan covenants, and provide additional details on relevant regulation.

Naturally, the formalization of the constraints ignores some differences between asset-based loans and loans subject to earnings covenants in reality. For example,

two.

[34] With the term "book value", I do not refer to the value at historical costs, but to the value that does not take into account market price variation arising from adjustment costs.

[35] Valuation by multiples is a common practice for assessing various types of assets and investment opportunities, see Damodaran (2012) for a textbook treatment.

[36] Greenwald (2017) rationalizes borrowing constraints for household mortgages along similar lines.

while collateral is pledged upon origination and may be seized in the case of default, covenants can in principle be exercised at any point during the lifetime of a loan. I abstract from these differences on two grounds. First, the fact that only the specific variable entering the right hand side of the debt limit is different between (1.8) and (1.9) allows for transparency in characterizing the implied differences in business cycle dynamics. Second, the Dealscan data shows that the maturity of corporate debt is relatively short, in particular compared to household debt, and that the relation between lenders and borrowers in the commercial loan market entails repeated interaction both in relation to covenant assessment and in relation to collateral.[37] The latter observation also justifies the simplification that both borrowing constraints affect one-period debt, which abstracts from considerations regarding maturity choice.[38]

Firm's maximization problem. The objective of the firm is to maximize the expected discounted stream of the dividends paid to its owner, that is, its maximization program is

$$\max \ \mathbb{E}_0 \sum_{t=0}^{\infty} \Lambda_t d_t \tag{1.11}$$

subject to (1.1), (1.2), (1.3), (1.4), (1.6), and either of the borrowing constraints (1.8) or (1.9). The term Λ_t in the objective function is the firm owner's stochastic discount factor between periods 0 and t. The firm's optimality conditions are shown in Appendix 1.7.4.

Household, government and equilibrium

Details on the household problem, the government and the definition of the equilibrium can be found in Appendix 1.7.4. The household consumes the good produced by the firm and supplies labor. She does not receive the tax rebate on debt and therefore becomes the saver in equilibrium. The government runs a balanced budget in every period.

It is worth noting that the model presented here does not feature a *labor wedge*, as the marginal product of labor (MPN) equals marginal rate of substitution (MRS) between consumption and leisure in equilibrium. I have explored extensions of the model in which the firm requires working capital loans to pre-finance expenditures. Since the results in this section are about *qualitatively* different borrowing dynamics arising from the alternative constraints, I stick to the simplest version in which $MRS = MPN$.[39]

[37] In the Dealscan data, the average (median) maturity of loans in 52 (60) months, and the value-weighted share of loans that refinance a previous loan is 83%.

[38] For a general equilibrium treatment of long-term debt, see for example Gomes, Jermann, and Schmid (2016). Cao and Nie (2017) provide a study of the role of market incompleteness implied by the non state-contingency of debt that is typically assumed alongside borrowing constraints.

[39] The literature, in particular Jermann and Quadrini (2012), has advocated the working capital formulation as a way to introduce an interaction between the labor wedge and financial frictions as an important amplification mechanism that delivers quantitatively more elevated responses to shocks. See also Chari, Kehoe, and McGrattan (2007) for a general discussion of the labor wedge

1.3.2 Model parameterization and specification

The stochastic processes underlying the exogenous disturbances follow autoregressive processes of order one in logs. See Appendix 1.7.4 for details. I specify the investment adjustment costs as a quadratic function that satisfies the functional form assumptions introduced by Christiano, Eichenbaum, and Evans (2005) and has been used in various subsequent papers on US business cycles, that is,

$$\Phi_t\left(\frac{i_t}{i_{t-1}}\right) = \frac{\phi_t}{2}\left(\frac{i_t}{i_{t-1}} - 1\right)^2. \tag{1.12}$$

This specification gives a steady state market value of capital of 1.[40,41] Furthermore, in steady state, $\Phi''(1) = \overline{\phi}$.

Panel (a) of Table 1.2 summarizes the values I set for the structural parameters of the model. Most parameter values are standard in business cycle research for the US case or match standard moments in US macroeconomic data. I set $\overline{\phi} = 4$ in line with Smets and Wouters (2007). To parameterize β, I calculate the average interest rate faced by firms in the Dealscan data base.[42] Panels (b) and (c) of the table show the calibration of the parameters that are related to the alternative borrowing constraints (1.8) and (1.9). In this part of the paper, I investigate model dynamics using the simplification that either one or the other constraint is faced by the firm. To do this, I exploit the fact that the model nests restricted versions in which only a collateral or only an earnings-based constraint are present. Each constraint can be shut off by parameterizing $\theta_j = \tau_j = 0$, for $j \in \{k, \pi\}$ and $\forall t$. In this case debt type j is in zero net supply and the other constraint binds at all times.[43]

I set the tax advantage of debt τ_j to 0.35 following Hennessy and Whited (2005). Regarding the tightness parameters of the constraints I proceed as follows. Using the Dealscan data I calculate the dollar-weighted mean covenant value of the debt-to-EBITDA covenant, the empirical counterpart of my earnings-based constraint. This gives a value of $\theta_\pi = 4.6$ (see Table 1.1). As this value is for annualized EBITDA and my model is quarterly, I divide by four. I then set the tightness of the collateral component to that value which achieves the exact same steady state debt level, which

in business cycle models.

[40]For the result presented in this section the specific form of adjustment costs is not crucial. For example, the conclusions drawn from the results are the same with adjustment costs in capital rather than investment. I choose this specification mainly to be consistent with Section 1.5.

[41]In order to study the adjustment cost shock to ϕ_t I introduce a minor modification to (1.12) in which steady state adjustment costs exceed zero by an arbitrarily small magnitude. This is done in order to be able to compute IRFs to this shock as deviations from the nonstochastic steady state. See more in Appendix 1.7.4.

[42]In particular I use the sum of the "All-in spread drawn", and add the 12-month LIBOR rate. I then calculate the mean over loan deals which feature either collateral, earnings-related covenants, or both.

[43]Throughout my analysis I focus on binding borrowing constraints. This assumes that shocks are small enough in magnitude to keep the Lagrange multiplier on the constraint positive, that is, $\mu_{jt} > 0$, $j \in \{k, \pi\}$, $\forall t$. Modifying my model to feature occasionally binding constraints would be relatively straightforward. This would make it feasible to also study possible switching effects between different types of borrowing constraints over the business cycle, similar to what Greenwald (2017) and Ingholt (2018) emphasize for the case of household mortgages. I leave an extension in this direction for future work.

Table 1.2: MODEL PARAMETERIZATIONS

Parameter	Value	Details on parameterization
		(a) Structural parameters
α	0.33	Capital share of output of 1/3
δ	0.025	Depreciation rate of 2.5% per quarter
$\overline{\phi}$	4	Prior of Smets and Wouters (2007)
β	0.9752	Steady state annualized interest rate of 6.6%*
χ	1.87	Target $n = 0.3$ in steady state
ψ	0.46	Jermann and Quadrini (2012)
		(b) Model with earnings-based constraint only
θ_k	0	Shut off collateralized borrowing
τ_k	0	Shut off collateralized borrowing
θ_π	4.6/4	Average value of debt-to-EBITDA covenants*
τ_π	0.35	Following Hennessy and Whited (2005)
		(c) Model with collateral constraint only
θ_k	0.0485	Same steady state debt as Panel (b)
τ_k	0.35	Following Hennessy and Whited (2005)
θ_π	0	Shut off earnings-based borrowing
τ_π	0	Shut off earnings-based borrowing

Note: Panel (a) describes the parameterization of the structural parameters which are the same independent of which type of constraint is specified to feature in the model. Panels (b) and (c) present the parameterizations that achieve that the firm faces either one or the constraint. * indicates parameter values that are calculated directly from the micro data, using ThomsonReuters Dealscan.

results in $\theta_k = 0.0485$. It should be emphasized that the results shown in the stylized model environment of this section are robust to variations in these parameter values. In particular, as the model is linearized and I focus on qualitative predictions, the results are not sensitive to varying the θ parameters across a range of values.

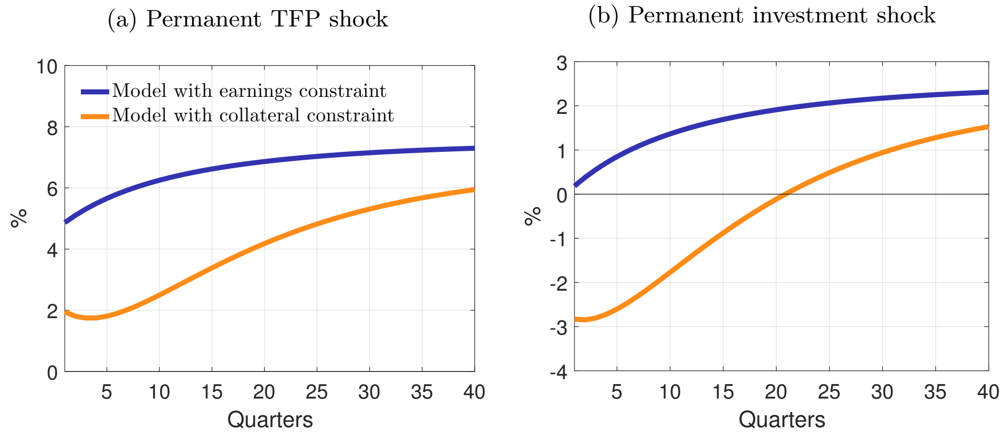
1.3.3 Dynamics implied by earnings-based vs. collateral constraints

Figure 1.2 plots the IRFs of firm debt to a positive TFP shock and a positive investment shock. Both shocks are permanent.[44] The dark blue lines correspond to the model in which firms face the earnings-based constraint (parameterization shown in Panel (b) of Table 1.2), while the light orange lines are generated in a model where the collateral constraint is present (see Panel (c)). The figure shows that while the responses of firm debt to the TFP shock are positive under both alternative borrowing constraints, the sign of the responses for the investment shock flip between one

[44] I show the results for permanent shocks since the SVAR methodology in Section 1.4 will allow me to identify permanent rather than transitory shocks in the data. The qualitative conclusions regarding the sign of the responses on impact are similar with transitory persistent shocks. See also Figure 1.3 further below.

and the other parameterization, implying the opposite comovement of debt with the shock. In other words, different conclusions about the dynamics of firm borrowing are drawn depending on how the borrowing friction of the firm is formulated.

Figure 1.2: MODEL IRFS OF FIRM DEBT UNDER DIFFERENT BORROWING CONSTRAINTS



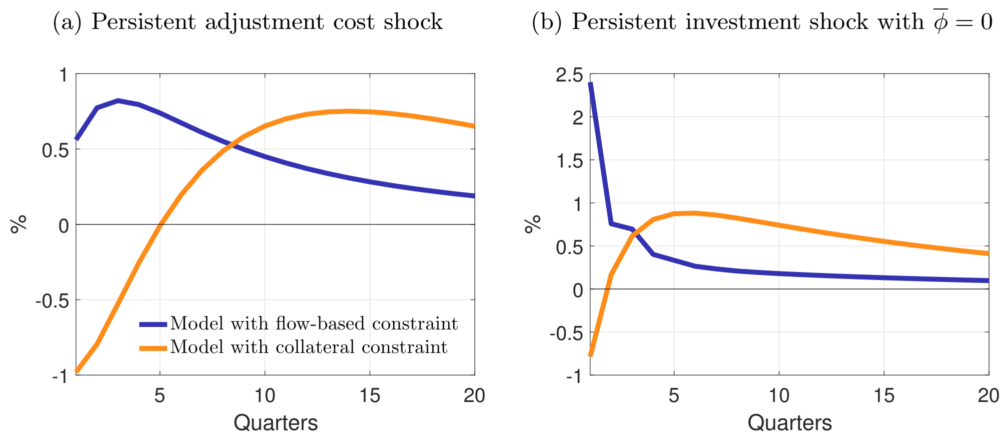
Note: The figure displays model IRFs of firm debt to different shocks, under two alternative calibrations in which only the earnings-based constraint (dark blue line) or only the collateral constraint (light orange line) is present. Panel (a) show the debt IRF to a positive TFP shock and Panel (b) to a positive investment shock. The parameters to generate these IRFs are shown in Table 1.2. I set $\rho_z = \rho_v = 1$ (the shocks are permanent) and $\sigma_z = \sigma_v = 0.05$. The figure highlights that the responses of debt to investment shocks have a different sign under the alternative borrowing constraints.

The intuition behind these dynamics is as follows. The TFP shock raises both the firm's earnings as well as the market value of capital, supporting more debt under both constraints. While the magnitudes differ, the sign of the debt responses to this shock are therefore the same under the alternative constraints. This is different for the investment shock, which leads to higher efficiency in the economy's investment margin. This induces investment and stronger economic activity accompanied by growing earnings. However, the shock reduces the relative value of capital in consumption units. This means that if the firm faces a collateral constraint, it needs to reduce its debt level, while it is able to borrow more in the face of an earnings-based constraint. The responses to this shock thus imply sharply different debt dynamics depending on the relevant borrowing constraint. These differences will provide the testing ground for my empirical analysis in Section 1.4.

As an illustration of the mechanism, think about an airline and imagine a shock – an exogenous technological innovation – that makes the production of airplanes cheaper, which lowers their price relative to other goods in equilibrium. The implication of this shock for borrowing differs sharply depending on the relevant constraint. If airplanes serve as collateral, their falling relative value tightens the borrowing constraint. By contrast, the earnings-based borrowing constraint is relaxed as cheaper airplanes increase the airline's profitability.

As discussed above, when the production of capital and investment goods are not disaggregated into separate sectors, a shock to v_t can be thought of as both an investment-specific technology (IST) and a marginal efficiency of investment (MEI)

Figure 1.3: MODEL IRFS OF FIRM DEBT: ADDITIONAL INVESTMENT MARGIN SHOCKS



Note: The figure displays IRFs of firm debt to additional investment margin shocks generated from the model, under the two alternative calibrations in which only the earnings-based constraint (dark blue line) or only the collateral constraint (light orange line) is present. Panel (a) plots the IRFs to a negative adjustment costs shock with $\rho_\phi = 0.5$ and $\sigma_\phi = 1$. Panel (b) repeats the investment shock IRFs from Figure 1.2 as a transitory but persistent shock ($\rho_v = 0.5$ and $\sigma_v = 0.05$) and without investment adjustment costs ($\overline{\phi} = 0$). The different signs of the IRFs across models show that the proposed mechanism is broad enough to carry through to different types of shock to the investment margin.

shock.[45] At a later stage in the analysis, for the purpose of the empirical verification of the mechanism in Section 1.4, I will narrow down the interpretation of v_t to capture IST. This allows me to establish a mapping of v_t to the data. At this stage, in terms of the main message behind the results, the distinction between these refined concepts is not of first order importance. In fact, the proposed mechanism has a broad interpretation which carries through to other shocks that affect the economy's investment margin. To demonstrate this, Figure 1.3 plots two more sets of IRFs. In Panel (a), the IRFs to a negative persistent adjustment cost shock for the two model versions are plotted. This is another disturbance that distorts the economy's investment margin. It is evident that this shock also results in a different sign of the debt responses on impact depending on which constraint is at play. In Panel (b), I repeat the IRFs to the investment shock from Figure 1.2, but shut off any adjustment costs and specify the shock as persistent rather than permanent. This corresponds to a setting in which there are no fluctuations in the price of capital other than through the exogenous disturbance itself. There is again a different sign of the impact response, with a positive debt response under the earnings constraint and a negative one when the collateral constraint is present. These additional responses highlight the broad scope of the key mechanism that the model delivers. Various types of disturbances that enter the investment margin give rise to different implications under the alternative constraints.

In Appendix 1.7.4 I repeat Figures 1.2 and 1.3 for a version of the earnings

[45]See the discussion below the introduction of equation (1.3) in Section 1.3.1. More details on this distinction is contained in Justiniano, Primiceri, and Tambalotti (2011) and Schmitt-Grohe and Uribe (2012).

constraint in which current and three lags of earnings enter the constraint. This is based on the idea that covenants are often evaluated based on a 4-quarter trailing average of the indicator. The results for this specification are similar to the ones shown in the above figures. The shape of the IRFs changes due to the fact that current earnings will affect the borrowing ability in future periods. In particular, there is a delayed and hump-shaped response under this version of the earnings-based constraint, but the signs of the responses remain unchanged.

Note that in deriving testable model predictions I focus primarily on the IRFs of debt. I turn to selected additional variables in the next subsection, and show the IRFs of remaining model variables in Appendix 1.7.4. The appeal of this strategy is that debt dynamics are tied very directly to the alternative constraint formulation and are not driven by further modeling choices on the structure in which they operate. Interestingly, in a prototype neoclassical setting under standard calibrations, debt constraints themselves typically do not have strong effects on the model's overall dynamics. Cordoba and Ripoll (2004) provide a detailed exploration of this insight.[46,47] I therefore show the responses of other variables only insofar as they help me to understand the different debt dynamics across parameterizations of the model. In Section 1.5 of the paper I do consider the dynamics of typical macroeconomic variables of interest.

In summary, the results highlight the different *qualitative* conclusions that can be drawn about the dynamics of debt depending on the type of borrowing constraint. In the next subsection I provide a more in-depth characterization of these results with an explicit focus on the theoretical link between asset values, earnings and other flow variables. After this additional discussion I turn to verifying the model predictions in US data in Section 1.4.

1.3.4 Discussion: borrowing against flow vs. stock variables

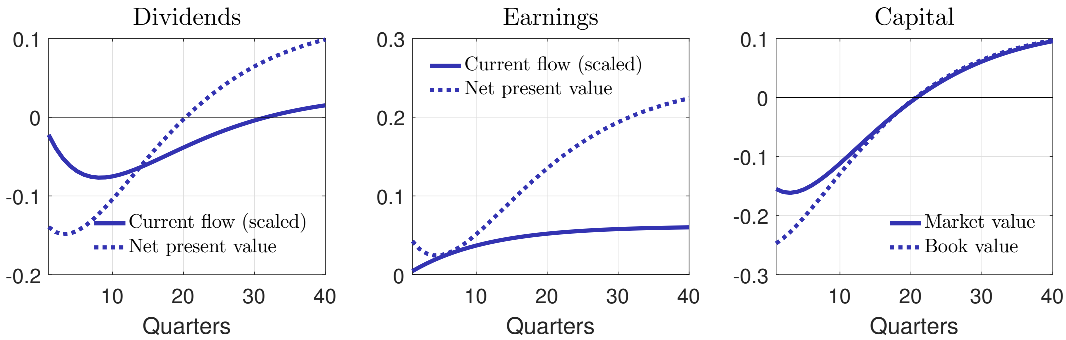
The analysis highlights the differences between two variables limiting the access to debt for firms: earnings and the value of capital, a *flow variable* and a *stock variable*, respectively. To further characterize the results, this section analyzes to what extent this difference has an influence on the differential responses to investment shocks. From a theoretical point of view, the market value of an asset corresponds to the net present value (NPV) of future flows that can be derived from that asset. In the context of a firm, its market value is equal to the flows that the firm generates for its owner. Several observations can be made on how the firm's market value and the

[46] A similar discussion is provided by Kocherlakota (2000), who shows that the amplification generated by credit constraints in a small open economy setting is sensitive to the quantitative specification of the underlying model structure, in particular factor shares. In related work, Fuerst (1995) shows that agency frictions in the spirit of Bernanke and Gertler (1989) also add little amplification in a basic business cycle model.

[47] As shown in Appendix 1.7.4, apart from the debt IRFs the model behaves extremely similar under the two constraints. This is different for example when raising the value of ψ or when choosing a working capital formulation, as discussed after the introduction of the firm's maximization problem above. The predictions on the qualitative dynamics of total firm debt, however, are not altered by these modifications of the model.

flows to the owner relate to the specific variables constraining debt in my model.

Figure 1.4: IRFS OF DIFFERENT FLOW AND ASSET VALUES TO INVESTMENT SHOCK



Note: The figure displays model IRFs of selected variables to a permanent investment shock, generated from a version of the model without any debt. This is intended to highlight the relation between alternative flows and asset values which may affect the right hand side of potential borrowing constraints. The unit of the IRFs is in levels of consumption units in the model (earnings and dividend flows are additionally scaled by 10). The net present values (NPVs) are recursively computed in the model using the household's stochastic discount factor.

First, in the equilibrium of the model, the market value of the firm corresponds to the NPV of *dividend* flows. That is, the firm's overall value is the infinite stream of d_t, discounted at the stochastic discount factor of the household $SDF_{t,t+1} \equiv \frac{\Lambda_{t+1}}{\Lambda_t} = \frac{\beta u_{c_{t+1}}}{u_{c_t}}$. We can define the market value of the firm recursively as $V_{d,t} = d_t + \mathbb{E}_t(SDF_{t,t+1}V_{d,t+1})$. Importantly, this value of flows is different both from the current earnings flow π_t as well as from the NPV of earnings flows, which can also be recursively defined as $V_{\pi,t} = \pi_t + \mathbb{E}_t(SDF_{t,t+1}V_{\pi,t+1})$. Second, in a neoclassical production economy, the market value of a firm is proportional to the capital it owns if specific conditions on technology are satisfied (see Hayashi, 1982): if technology is constant returns to scale and adjustment costs are homogeneous of degree 1 in k, it is the case that $V_{d,t} = Q_t k_{t-1}$. In this context Q_t is known as "Tobin's Q". As a consequence of the two observations, if the conditions of Hayashi (1982) hold, the collateral constraint is equivalent to a constraint in which the firm's overall market value serves as collateral. In turn, this constraint would have an equivalent flow-related analogue, if the flows entering the flow constraint were all discounted future dividend flows. In this case, the two borrowing limits would be equivalent.

In light of these theoretical insights, we can see that the earnings-based borrowing constraint (1.8) and the collateral constraint (1.9) are not equivalent for two reasons. First, they differ in terms of the *flow definition*. The earnings-based constraint features earnings rather than dividends. Second, they differ in terms of the *flow timing*. The earnings-based constraint features a current flow variable rather than the NPV of all current and future flows.[48] In the model, I can check directly which of these two differences drives the results in Figure 1.2, by comparing the responses

[48] As shown in the previous section, I have explored sensitivity of the results to generalizations of the earnings-based constraint where lagged or one period ahead expected earnings can enter. These versions are still very different from the NPV, so the arguments made here again carry through.

of d_t, $V_{d,t}$, π_t, $V_{\pi,t}$ and $Q_t k_{t-1}$ to the investment shock. Figure 1.4 displays these IRFs in a model without borrowing constraints. This is essentially a comparison of different variables that could potentially appear on the right hand side of a borrowing limit. The figure shows that both current earnings as well as the NPV of earnings rise in response to the shock. This means that with an earnings constraint additional debt could be issued in response to the investment shock and that timing of earnings by itself is not key in this case. However dividends as well as the NPV of dividends, which equal the firm value and the value of the capital stock under the Hayashi conditions, is reduced.[49] This leads to the counterfactual debt response with the collateral constraint. Hence, for the investment shock the difference in the debt response is driven by the flow *definition*. The results in Panel (b) of Figure 1.2 arise not because debt is constrained by a flow instead of by an asset value *per se*, but by the specific variable that defines this flow, current earnings.

1.4 Verifying the model predictions for investment shocks

This section empirically verifies the model predictions implied by the alternative borrowing constraints. First, I investigate which of the two borrowing limits, earnings-based or collateral constraint, is consistent with comovements we observe in US macroeconomic data. Second, I examine if the dynamics in the data are in line with the specific mechanism through which the constraints on the firm operate in the model. I resort to analyzing both aggregate and firm-level data, using an SVAR (Section 1.4.1) and a panel regression framework that allows for heterogeneous responses to shocks (Section 1.4.3).

The empirical analysis focuses on the structural shock that has given different qualitative predictions in the model: the investment shock. As explained in Section 1.3, the disturbance v_t can capture both shocks to investment-specific technology (IST) as well as to marginal investment efficiency (MEI). The former type is directly tied to a readily available empirical counterpart, the inverse relative price of investment goods.[50] Observable time series of this price have been exploited by previous research to identify IST shocks. I build on this work to study the conditional dynamics of US data with a focus on the debt responses to investment shocks. That is, while the interpretation of the model mechanism can be broadly applied to different shocks to the investment margin, for the purposes of verifying the predictions empirically, I interpret v_t as a specific type of investment shock, a shock to IST.[51]

[49] Under the functional form of investment adjustment costs chosen in (1.12), the Hayashi conditions are not satisfied in the model (see also Jaimovich and Rebelo, 2009). However in the calibration the numerical difference between NPV of dividends and the market value of capital is very small, as can be seen from the similarity bewteen the dashed line in the left chart and the solid line in the right chart of Figure 1.4.

[50] In a subset of the loan-level data from Dealscan, it is possible to directly observe the *type* of collateral that is used in loan facilities. After excluding non-informative categories such as “Other” and “Unknown”, the category “Property & equipment” is the largest one, three times as large as “Real Estate”, both in terms of the number of facilities and the dollar volume. See Table 1.7 in the Appendix.

[51] Justiniano, Primiceri, and Tambalotti (2011) emphasize that MEI shocks are more important

1.4.1 SVAR on aggregate US data

I specify an SVAR framework to estimate the impact of IST shocks on the US economy as a whole. The system includes variables that allow me to distinguish between dynamics that are supportive of either the earnings-based constraint or the collateral constraint: debt, earnings and capital. I use two different identification schemes. First, following the literature on technology shocks in SVARs, I identify IST shocks using long-run restrictions building on the work of Fisher (2006).[52] Second, I use medium-run restrictions following Barsky and Sims (2012), and Francis, Owyang, Roush, and DiCecio (2014).[53] I apply both identification methods to US postwar data. In addition, I set up a Monte Carlo experiment in which I repeatedly run the SVAR model on data that I generate directly from the model, in order to check the SVAR's ability to distinguish between the alternative borrowing constraints.

SVAR setting and identifying assumptions

I begin by formally introducing the general setting that encompasses both identification methods. Consider the n-dimensional vector of macroeconomic time series Y_t, which is specified to follow

$$B_0 Y_t = B_1 Y_{t-1} + ... + B_p Y_{t-p} + u_t, \tag{1.13}$$

where the vector u_t denotes the structural shocks with covariance matrix $\Omega_u = I_n$. The model can be rewritten in its $MA(\infty)$-representation as

$$Y_t = B(L)^{-1} u_t, \tag{1.14}$$

where L denotes the lag operator. The structural shocks u_t are not identified unless additional restrictions are imposed on the parameters of the system.

Identification using long-run restrictions. The idea behind long-run restrictions is to impose identifying assumptions on the long-run multiplier $B(1)^{-1} = [B_0 - B_1 - ... - B_p]^{-1}$. Following the seminal study of Fisher (2006), I use as the first three variables the log difference of the relative price of investment, the log difference in output per hour, and the log of hours. The idea is to identify two shocks, using a recursive scheme on $B(1)^{-1}$: the long-run level of the first variable is only affected by the first shock, and the long-run level of second variable is only affected by the first and the second shock. The first shock has the interpretation of

than IST shocks for US business cycles. MEI shocks, however, are not directly identifiable the same way that IST shocks are. It will turn out that the IST shock I identify is reasonably important in terms of the historical variance decomposition of debt implied by the SVAR.

[52] Long-run restrictions are the most common way to identify technology shocks in SVARs. Blanchard and Quah (1989) and Gali (1999) are early contributions which focus exclusively on TFP. Fisher (2006) and various subsequent papers also estimate the effect of IST shocks. A recent example is Ben Zeev and Khan (2015).

[53] Ramey (2016) provides a useful summary of the literature on both long-run and medium-run restrictions.

investment-specific technological change, as the relative price of investment is only affected by this shock in the long run. The second shock represents a concept akin to a TFP shock, as it is the only driver that affects, other than IST, the economy's labor productivity in the long run.[54] It is important to highlight that these restrictions are satisfied in the model of Section 1.3. For the purpose of this paper, I view the identification of the TFP shock as a by-product and mainly present the model results for the IST shock, as the latter shock implies sharply contrasting predictions under the alternative borrowing constraints.

As I only identify two shocks and leave the remaining rows of $B(1)^{-1}$ unrestricted, I can add further variables to the system, for which the ordering becomes irrelevant to the identification of IST and TFP. The additional variables are the log difference in aggregate business earnings, the log difference in the relative value of the capital stock and the log difference in business sector debt. In particular the inclusion of debt is key, as I have shown that in the model this variable responds with a different sign to investment shocks depending on the borrowing constraint component that is present. Together this gives, in line with the notation of the model,

$$Y_t = [dlog(p_{kt}) \;\; dlog(y_t/n_t) \;\; log(n_t) \;\; dlog(\pi_t) \;\; dlog(p_{kt}k_t) \;\; dlog(b_t)]'. \tag{1.15}$$

p_{kt} is the relative price of investment, which corresponds to v_t^{-1} if v_t captures IST.

Identification using medium-run restrictions. The idea behind medium-run restrictions is to identify a shock such that its forecast error variance decomposition (FEVD) share for a selected variable at a specific finite horizon h is maximized. These restrictions have been introduced to overcome weaknesses of the long-run identification method, such as their small sample properties (for details see Faust and Leeper, 1997). Francis, Owyang, Roush, and DiCecio (2014), for example, identify a technology shock as the shock that maximizes the FEVD share of labor productivity at horizons of 2.5 to 20 years. Barsky and Sims (2012) implement a variant of this method where the shock maximizes the *sum* of the FEVD up to a specific horizon.[55] I follow the latter authors' variant of this identification scheme. Using the same vector of observables Y_t, I identify the IST shock as the shock that maximizes the cumulative FEVD share in the relative price of investment over varying horizons h. Again, I leave the remaining shocks unidentified.[56]

[54]This identification scheme implies that the first row of $B(1)^{-1}$ is composed of zeros apart from the first element and the second row is composed of zeros apart from the first two elements. Fisher (2006) also imposes the additional, overidentifying restriction that labor productivity responds in a fixed proportion to movements in the relative investment price. While this improves the precision of the estimates, I do not impose this restriction to remain as agnostic as possible.

[55]Earlier work on these types of restrictions includes Uhlig (2004) and Faust (1998). They are also applied in recent paper by Angeletos, Collard, and Dellas (2018).

[56]To implement the medium-horizon identification, I begin by estimating the VAR in reduced form and calculating an initial estimate of the B_0^{-1} matrix, based on a simple Cholesky decomposition. I then take an orthonormal rotation of this matrix such that the identifying restriction is satisfied. In practice, this means that I run a constrained optimization routine over $n \times n$ matrices D, in which

Data used for SVAR analysis

I use data from the US National Income and Product Accounts (NIPA), and the US Financial Accounts (Flow of Funds) for the total nonfinancial business sector. Details can be found in Appendix 1.7.1. To compute real variables I use nominal data which I deflate with the consumption deflator for nondurable goods and services. An important consideration lies in the choice of data for p_{kt}. Following the literature on IST shocks, I use the relative price of equipment investment.[57] I construct this relative price from NIPA data and use the Gordon-Violante-Cummins (GVC) investment price for robustness.[58] For debt I use the sum of loans and debt securities for the nonfinancial business sector and also consider these debt categories separately for robustness. As some of the variables display low frequency movements after log differencing, I detrend some of the series before estimating the VAR.[59] I estimate the reduced form VAR using OLS, recover the IRFs from inverting (rotating) the relevant matrices under the identifying restrictions, and compute 68% error bands using bootstrap techniques.

1.4.2 SVAR results: aggregate responses to investment shocks

IRFs. The results on quarterly US data from 1952 to 2016 for $p = 4$ are shown in Figure 1.5. The figure presents the IRFs for a positive permanent IST shock identified based on its long-run impact on the relative price of capital. Section 1.7.5 in the appendix presents the analogous IRFs based on the medium-horizon identification scheme with $h = 20$ and $h = 40$, implying that IST is the main driver of the relative price of investment at a 5 and 10 year frequency, respectively. For both identification methods the figure shows a positive response of debt, which is in line with the model predictions for the earnings-based constraint but not for the collateral constraint. In line with the dynamics of the model, the rise in debt is accompanied by growing earnings and a fall in the value of capital.

To best interpret these results, suppose the model introduced in Section 1.3 approximates well the data-generating process behind the time series used in the SVAR. Given the positive debt response to the IST shock in Figure 1.5, the IRFs are supportive of the version of the model that features an earnings-based constraint and not a collateral constraint. The dynamics in US data, conditional on identified

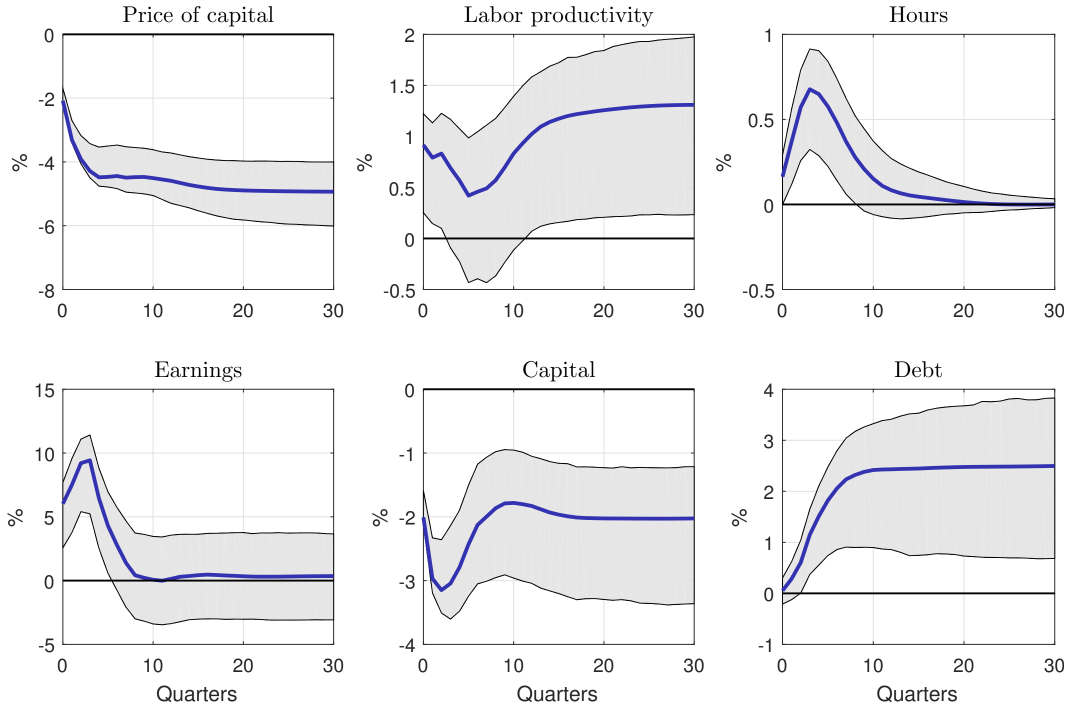
I calculate the FEVD implied by DB_0^{-1}, and the objective is to maximise the FEVD share of the first shock in the first variable. The constraint of the routine is that $D'D = I_n$ must be satisfied.

[57] Among different relative investment price categories, the equipment price is the one with a clear downward trend in US data based on which the IST shock can be identified. The deflator for total investment and for structures do not display as strong trends. Interestingly, the Dealscan data set provides information on the type of collateral used in loan contracts: "property and equipment" is the top category of specific assets pledged as collateral, which further supports this choice. See Table 1.7 in the Appendix.

[58] Appendix 1.7.1 contains a figure comparing the two alternative series. I also run a unit root test to confirm that both series are nonstationary in levels and stationary after first-differencing, as required by the identification scheme of the SVAR. See also Gali (1996) for details.

[59] Blanchard and Quah (1989) provide a related discussion. The detrending mostly increases the precision of the estimates but has little influence on the shape of the estimated IRFs.

Figure 1.5: SVAR IRFS TO POSITIVE IST SHOCK IDENTIFIED WITH LR RESTRICTIONS



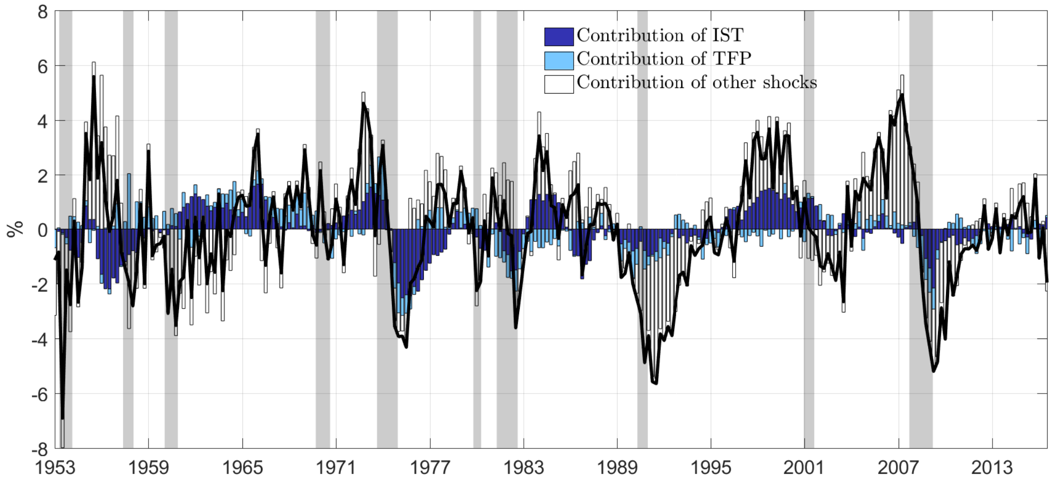
Note: The figure displays the IRFs to an investment-specific shock identified from an estimated SVAR model using US data. The identification scheme relies on long-run restrictions following Fisher (2006). The responses are shown for all six variables included in the system, in percent. The unit of the shock is one standard deviation. The sample period used for estimation is 1952:Q2 to 2016:Q4. 68% error bands are calculated using bootstrap techniques. The figure shows a positive response of debt to an investment shock, which is in line with the predictions arising from a earnings-based borrowing constraint in the theoretical macro model.

shocks, thus lend support to the importance of earnings-based borrowing for debt dynamics in US business cycles.[60]

Historical variance decomposition. My empirical strategy only relies on the sign of the responses, and conceptually it does not require the importance of the shock in a variance decomposition sense to be large. However, if the shock is an important driver of macroeconomic dynamics, this means that the relevance of the earnings-based borrowing constraint should have important effects also on the unconditional dynamics of the macro data. Figure 1.6 shows the historical decomposition of debt using the long-run identification method. The solid black line plots the actual data of the cyclical component of debt, and the colored bars represent the contribution at each point in time of the different shocks (IST, TFP and other). It can be seen that IST shocks have played a marked role in different episodes of the postwar US business cycle. For example, consistent with the narrative around the tech boom,

[60] Appendix 1.7.5 presents the IRFs to the TFP shock. Consistent with the model, this shock has an expansionary effect, raising the variables in the system, despite hours. Debt also rises (albeit not significantly). For TFP shocks, however, it did not make a big difference which constraint is relevant to begin with, so the empirical verification of the specification of the borrowing constraint relies on the responses to the IST shock.

Figure 1.6: SVAR: HISTORICAL VARIANCE DECOMPOSITION OF FIRM DEBT



Note: The figure shows the historical variance decomposition of firm debt estimated by the SVAR model identified with long-run restrictions. The black line is the percent deviation from trend of debt liabilities (loans and debt securities) of the nonfinancial business sector, taken from the US financial accounts. The bars indicate the contribution of different structural shocks to the variance of debt as estimated by the SVAR model. The dark blue bars represent investment shocks, the light blue ones TFP shocks, and the contribution of shocks that remain unidentified are shown by the white bars. Shaded areas indicate NBER recessions. The decomposition is calculated following Kilian and Lütkepohl (2017).

the 1990s expansion was strongly driven by IST. The boom and bust of the 2000s, on the other hand, was less influenced by IST according to the SVAR model. Appendix 1.7.5 provides the historical decompositions of the remaining variables in the system.

Monte Carlo simulations. To verify the ability of the SVAR methodology to distinguish between different borrowing constraints, I set up a Monte Carlo experiment in which I estimate the SVAR on simulated data generated from the model in Section 1.3. Specifically, I repeatedly create two types of data samples, each generated from one of borrowing constraint specification (Panel (b) vs. Panel (c) in Table 1.2). I do so by randomly drawing TFP, IST and additional shocks and then simulating the time series in (1.15) from the policy rules of the model. For each sample type I generate 10,000 repetitions and run the SVAR identified with long-run restrictions on each of them. The results, shown in Appendix 1.7.5, are reassuring. For example, the negative debt response generated from a collateral constraint model is fully contained in the 68% confidence set across Monte Carlo repetitions.

Robustness checks. I explore robustness of the SVAR results along several dimensions. First, following Fisher (2006), I split the sample in the early 1980's to account for the change in the trend exhibited by the relative price of investment. In the first part of the sample the shapes of the IRFs are preserved, while the bands get wider. In the second part, the debt response to IST is again positive and significant, but more hump-shaped rather than settling at a permanent level. Second, I construct the business debt time series separately for loans and debt securities.

This split reveals that the debt IRF in Figure 1.5 is mainly driven by loan dynamics, while the response for debt securities is noisy, and even negative for the first three quarters. Third, similar to many papers in the IST literature, I use the Gordon-Violante-Cummins (GVC) relative equipment price series as opposed to the relative NIPA deflator as an alternative measure of the relative price of equipment.[61] The results are very similar to the ones obtained using NIPA data. Finally, as the data on investment deflator dynamics is subject to a few large spikes, I also adjust this data for outliers as a robustness check. The IRFs get smaller in magnitude, but their shapes and statistical significance is preserved.

1.4.3 Panel projections in firm-level US data

The SVAR results indicate that the earnings-based constraint is relevant for debt fluctuations in the aggregate economy. The responses of aggregate earnings and capital are consistent with the model mechanism. In this subsection, I exploit micro-level information on how firm borrowing is restricted to verify the proposed mechanism more directly. I merge the Dealscan data set used in Section 1.2 with balance sheet information from the Compustat Quarterly data base to obtain a firm panel that has information on earnings-based covenants and collateral as well as on rich firm characteristics. I regress firm-level borrowing on the macro investment shock obtained from the SVAR. I obtain average IRFs across all firms, as well as heterogeneous IRFs for different borrower types, allowing me to verify whether the suggested mechanism is plausible in generated debt dynamics at the firm level.[62]

Panel setting and assumptions

I estimate the IRF of borrowing of firm i at horizon h to the investment shock by running the linear regression

$$log(b_{i,t+h}) = \alpha_h + \beta_h \hat{u}_{IST,t} + \boldsymbol{\gamma} \boldsymbol{X_{i,t}} + \gamma t + \eta_{i,t+h} \tag{1.16}$$

and obtaining estimates of β_h, $h = 0, 1, 2, \ldots, H$. The right hand side variable $\hat{u}_{IST,t}$ denotes the time series of the identified exogenous investment shock from the SVAR model above. $\boldsymbol{X_{i,t}}$ is a vector of controls. t is a linear time trend. This regression is a panel version of the local projection method to estimate IRFs following Jordà (2005). Equation (1.16) gives an average IRF across *all* firms in the panel. My model predicts the response of debt to the investment shock in this regression to be positive under an earnings-based constraint ($\beta_h > 0$) and negative with a collateral constraint ($\beta_h < 0$).

[61] This series was originally constructed by Bob Gordon and extended by Cummins and Violante (2002). See also DiCecio (2009) for details. Appendix 1.7.1 contains a figure comparing the long-run trends and cyclical dynamics in the two relative investment prices.

[62] I also study the firm-level debt responses to a fall in the relative price of investment goods, using an IV strategy in which the estimated investment shock serves as the instrument. This is in the spirit of Jordà, Schularick, and Taylor (2017). Related studies on firm-level responses to macro shocks are Bahaj, Foulis, Pinter, and Surico (2018) and Cloyne, Ferreira, Froemel, and Surico (2018) who focus on monetary policy.

Given the information in the Dealscan data, I can interact the shock with dummies that capture whether a firm is subject to earnings-based covenants or uses collateralized loans, effectively obtaining heterogeneous IRFs across different borrower types. This allows me to verify the proposed theoretical mechanism more directly. Formally,

$$\begin{aligned} log(b_{i,t+h}) = \alpha_h + \beta_h \hat{u}_{IST,t} + \boldsymbol{\gamma X_{i,t}} \\ + \beta_h^{earn} \mathbb{1}_{i,t,earn} \times \hat{u}_{IST,t} + \alpha_h^{earn} \mathbb{1}_{i,t,earn} \\ + \beta_h^{coll} \mathbb{1}_{i,t,coll} \times \hat{u}_{IST,t} + \alpha_h^{coll} \mathbb{1}_{i,t,coll} + \gamma t + \eta_{i,t+h}, \end{aligned} \tag{1.17}$$

where $\mathbb{1}_{i,t,earn}$ and $\mathbb{1}_{i,t,coll}$ are dummy variables that capture whether the firm is subject to earnings-related covenants or uses collateral. Their data counterparts are discussed further below. The interactions with these dummy variables allow me to estimate heterogeneous IRFs for four different firm groups. In particular, the IRF of an "earnings only" ("collateral only") borrower at horizon h is given by the sum of the coefficients β_h and β_h^{earn} (β_h and β_h^{coll}). My theoretical mechanism predicts that $\beta_h + \beta_h^{earn} > 0$ and $\beta_h + \beta_h^{coll} < 0$.

An alternative version of (1.17) based on an IV strategy is provided in Appendix 1.7.6. The idea is to study the responses to a fall in the relative price of investment goods, instrumented by the exogenous investment shock, rather than considering the direct responses to the shock itself. The corresponding results are presented in the same appendix and are discussed below.

Data and specification used for panel regressions

The Dealscan-Compustat merge is enabled by a link file connecting the identifiers in the two data sets, which has been created by Michael Roberts and collaborators (see Chava and Roberts, 2008).[63] The final data set I use covers around 150,000 firm-quarter observations for more than 4,000 distinct firms from 1994 to 2015. $b_{i,t}$ is the quarterly level of debt liabilities from Compustat (calculated as the sum of the items 'dltq' and 'dlccq'). Consistent with the data treatment in the SVAR, I obtain a real series by deflating with the consumption deflator for nondurable goods and services. The firm-level classification into "earnings borrowers" and "collateral borrowers" based on the information in Dealscan is consistent with the aggregate shares I present in Figure 1.1. $\mathbb{1}_{i,t,earn}$ is equal to 1 if a given firm issues a loan with at least one earnings covenant. $\mathbb{1}_{i,t,coll}$ is equal to 1 if the debt issued by the firm is secured by specific assets (see the explanations provided in Section 1.2). As an alternative, I also construct a version of $\mathbb{1}_{i,t,coll}$ based on whether the firm uses a secured revolving line of credit.[64] Summary statistics for the full data sample and conditional on the first grouping are provided in Appendix 1.7.1.

[63] I am extremely grateful to these authors for publicly providing this link. More details about the construction of the merged data set can be found in Appendix 1.7.1.

[64] This follows Lian and Ma (2018), who point out that secured "revolvers" are typically asset-based.

I focus on the version of $\hat{u}_{IST,t}$ estimated using long-run restrictions in Section 1.4.1. To the extent that my identification in the SVAR is credible, this shock is a purely exogenous regressor, meaning that there are no endogeneity issues in (1.16). Clearly, however, the dummy interactions to generate heterogeneous responses in equation (1.17) are a cause for concern. There may be omitted variables that affect both the left hand side and the endogenous selection of borrowers into a particular type. I address this problem by controlling for omitted characteristics that may simultaneously be driving debt responses to investment shocks and selection into borrower types. Specifically, I use a specification with 3-digit industry-level fixed effects and firm size, as well as firm-level real sales growth to control for firm-specific cyclical conditions. In an alternative specification I also introduce firm-level fixed effects. In all versions of (1.16) and (1.17) that I estimate, I include one lag of the left hand side variable and a linear time trend to the regression. Furthermore, I add a control variable that is intended to capture macroeconomic shocks other than investment shocks, which I construct from the SVAR residuals.[65] I set $H = 12$, and keep the firm composition constant when expanding h, that is, I restrict the analysis to firms where debt information is available for the current and 12 quarters ahead.

It should be emphasized that while Compustat is an actual panel, the loan issuance information from Dealscan is "sparse" in the sense that firms only have an issuance that is captured in this data every other quarter. Many firms appear only a few times during the sample period, while their total debt liabilities are continuously recorded in Compustat. This has two consequences. First, using any Dealscan information at time t means that the sample to estimate (1.17) is restricted to those firms that have a loan issuance captured by the Dealscan data in period t, which reduces the sample relative to the one I can use to estimate (1.16). Second, this also implies that the sample used to estimate (1.17) is restricted to firms that issue any debt to begin with. While I address the endogenous selection into debt types, I cannot address the endogenous extensive margin selection into being a borrower.

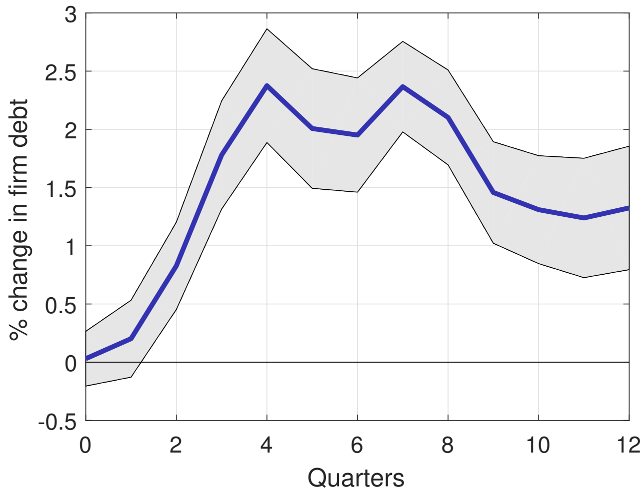
1.4.4 Firm-level results: heterogeneous responses to investment shocks

I first present the average debt response across all firm in the panel, that is, the estimates of β_h in (1.16) across horizons, together with the associated 90% bands. I cluster standard errors at the 3-digit industry level to allow for correlation in the residuals across firms within the same industry. In this regression I do not add any controls other than lags of the left-hand-side variable, a time trend and the exogenous shock itself. Figure 1.7 shows that the dynamic response of firm debt to an investment shock is positive, in line with the aggregate debt response in the SVAR, and consistent with the model in which the earnings-based constraint is the relevant debt limit. It matches the SVAR responses also in terms of the magnitude

[65] I use the reduced form residuals of the debt equation in (1.14) and orthogonalize them with respect to the structural IST shock. The resulting series captures innovations to aggregate debt that are unrelated to IST.

and persistence. This is reassuring, since Compustat-Dealscan firms are a specific subset of the total US nonfinancial business sector for which I use data in the SVAR.

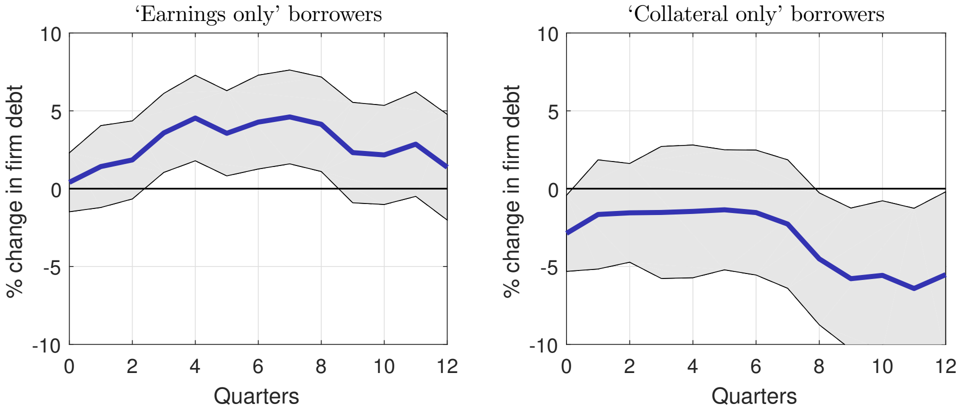
Figure 1.7: EMPIRICAL FIRM-LEVEL IRF OF DEBT TO AN INVESTMENT SHOCK



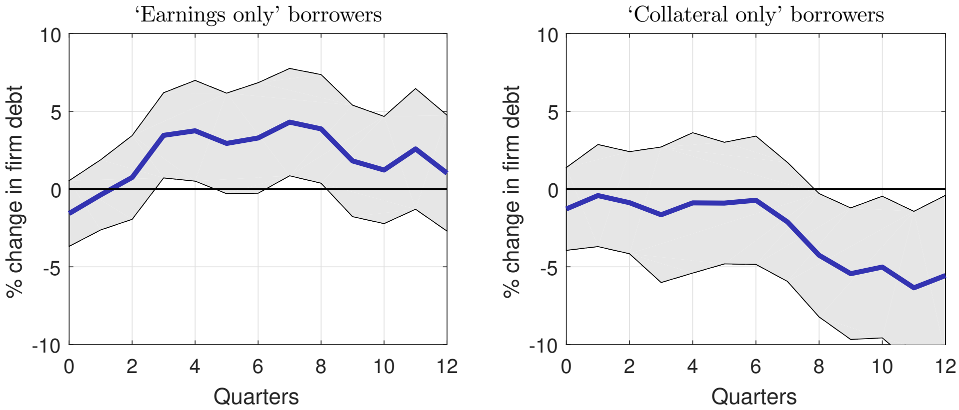
Note: The figure plots the average IRF of firm debt to a macro investment shock across individual firms, estimated using the method of Jordà (2005) in a panel data context, as formulated by equation (1.16). The macro shock has been identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information with balance sheet variables from the Compustat quarterly data base. The IRF is shown in percent. 90% bands are calculated using standard errors clustered at the 3-digit industry level. The figure shows that the debt IRF matches the one of aggregate debt in SVAR model and is in line with the predictions arising from an earnings-based borrowing constraint in the theoretical macro model.

Figure 1.8: FIRM-LEVEL IRFS OF DEBT TO IST SHOCK FOR DIFFERENT BORROWER TYPES

(a) Using collateral classification based on specific assets



(b) Using collateral classification based on secured revolvers



Note: The figure displays average IRFs of firm borrowing within different firm groups, estimated using the method of Jordà (2005) in a panel context, as formulated by equation (1.17). In both panels, the debt IRF for borrowers with earnings covenants and no collateral (left) and borrowers without earnings covenants but with collateral (right) are plotted. The results are based on a specification with detailed firm-level controls (3-digit industry fixed effects, size as measured by number of employees, growth of real sales and other macroeconomic shocks). Panel (a) uses the collateral classification based on whether a loan is backed by specific assets or not (see details in Section 1.2). Panel (b) uses an alternative grouping where secured revolvers are categorized as collateralized debt (see Lian and Ma, 2018). The investment shock is identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information, with balance sheet variables from the Compustat quarterly data base. 90% bands are calculated using standard errors clustered at the 3-digit industry level. The IRFs shown in the figure are consistent with the model's prediction of a positive debt response under an earnings-based constraint and a negative one under a collateral constraint. A formal test rejects the null hypothesis of equal responses across the two firm types for various horizons, as shown in Table 1.12 in the Appendix. The results for alternative specifications, as well as the responses for the remaining two borrower types are given in Appendix 1.7.6.

The heterogeneous IRFs based on estimating equation (1.16) are presented in Figure 1.8. These results are based on a specification with 3-digit industry fixed effects, size as measured by number of employees and growth of real sales. As discussed above, I also control for other macroeconomic shocks. Panel (a) shows the results where the classification of collateralized debt is based on whether a given firm's borrowing is secured with specific assets (see Section 1.2 for details). Panel (b) shows the results using the alternative classification of asset-based debt based on whether a firm uses secured revolvers (see Lian and Ma, 2018). Again, I plot 90% error bands based on standard errors that are clustered at the 3-digit industry level. The bands across all four figures are wider than in Figure 1.7 due to the lower number of observations when using $\mathbb{1}_{i,t,earn}$ and $\mathbb{1}_{i,t,coll}$ in the regression. Both panels of Figure 1.8 show that the IRF of debt to an investment shock is positive for firms that are subject to earnings-related loan covenants, but negative for firms that borrow against collateral. This confirms the key prediction of the model, as presented in Panel (b) of Figure 1.2. Interestingly, while the shape of the IRF for earnings-borrowers is similar to the model prediction – small on impact and then increasing persistently – the IRF of collateral borrowers differs from its model counterpart. Similar to the model prediction, the response on impact in Panel (a) is significantly negative. However it then rises and is again significantly negative after around 2 years. This may be due to the fact that one aspect of the theoretical mechanism I propose – the assumption that dynamics of new and already installed capital prices are the same – is not borne out by the data. Empirically, the relative price effect at the heart of my mechanism may not be strong enough to generate a negative effect for collateral borrowers that is as sizable as in the model. Reassuringly, the null hypothesis of an equal response across the two borrower is rejected over several horizons at the 5% level. This is not directly visible in Figure 1.8, but is formally presented in Table 1.12 in the Appendix.

Appendix 1.7.6 presents a host of additional results based on alternative variations of equation (1.17). First, I estimate the IRF to a fall in the relative price of equipment investment, instrumented by the investment shock (rather than the response to the investment shock directly). Second, the results for a specification based on firm fixed effects are presented. Finally, the appendix also shows the IRFs of Figure 1.8 for the two additional groups, which are firms subject to both earnings covenants and collateral, as well as firms that are subject to neither. Qualitatively, these results look very similar to the ones presented above. The exception is the firm fixed effect specification, where the debt response of collateral borrowers is flat and the response of earnings borrowers is positive in just one out of the two classifications.

1.4.5 Take-away: empirics in line with earnings-based constraint

The proposed model mechanism allows to distinguish between alternative borrowing constraints based on credit dynamics arising from investment shocks. In an economy in which firms are borrowing constrained, a boom driven by an expansionary shock which suppresses capital prices, debt levels rise in the presence of an earnings-based constraint, but not if capital serves as collateral. The empirical responses of debt to investment shocks in macroeconomic data, shown above, indicate that the relevant one for US aggregate corporate debt dynamics is such an earnings-based constraint. Moreover, heterogeneous firm-level responses are in line with the mechanism: earnings-based borrowers increase their debt liabilities in response to an aggregate investment shock, firms subject to collateral constraints do not.

1.5 Earnings-based borrowing in a quantitative macro model

This section extends the model of Section 1.3 to incorporate features of a New Keynesian quantitative macro model. Specifically, I add a number of shocks and frictions, such as price and wage rigidities to feature alongside borrowing constraints. I estimate the model on US time series to let the data speak about the importance of earnings-based borrowing relative to borrowing against collateral, and also relative to other frictions in the economy. This analysis goes beyond the focus on qualitative responses, which has guided my analysis above.

1.5.1 Setup of the quantitative model

The model is a New Keynesian DSGE model in the spirit of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). These models have become the workhorse model in central banks, perhaps due to their appealing philosophy: in order to gauge the overall effect of any macroeconomic policy change, this policy needs to be assessed net of other important forces that operate across parts of the economy.[66] For the purpose of adding borrowing constraints, I build on a variation of the Smets and Wouters (2007) model suggested by Jermann and Quadrini (2012).[67] The details of the model are provided in Appendix 1.7.7, in what follows I elaborate mainly on the borrowing constraints.

While in Section 1.3 I study alternative model calibrations in which either the earnings-based or the collateral constraint is binding, I now move to a formulation where both constraints are present simultaneously and where the estimation of the model can attribute a different relative importance to either constraint. Specifically, there is a continuum of firms which have access to a nominal risk-free bond that is constrained by a weighting between an earnings-based and a collateral component (in real terms). The interest rates paid on the debt is subject to a tax advantage of

[66] For recent discussions, see for example Galí (2018) and Christiano, Eichenbaum, and Trabandt (2018).

[67] The quantitative model of Jermann and Quadrini (2012) differs from Smets and Wouters (2007) in the following ways. Firms rather than households own and accumulate capital. Nominal rigidities arise because firms face Rotemberg price adjustment costs rather than Calvo pricing. The monetary policy maker targets output deviations from steady state rather than from the natural level. The exogenous disturbances do not feature moving average terms. Finally, firms have access to debt and receive a tax advantage on debt. I also add some corrections to the model that were suggested by Pfeifer (2016).

the type in equation (1.7). The constraint of firm i reads

$$\frac{b_{i,t}}{P_t(1+r_t)} \leq \omega\theta_{\pi,t}\pi_{i,t} + (1-\omega)\theta_{k,t}\mathbb{E}_t p_{kt+1}(1-\delta)k_{i,t}. \tag{1.18}$$

ω captures the relative weight on the earnings-based component in firm borrowing. The θ terms are subject to shocks to financial conditions. I choose the formulation of (1.18) as a reduced form way to capture that, in the aggregate, either constraint type will contribute to the dynamics to a certain degree. Using a weighting has the advantage that this degree can be captured directly by one single parameter without requiring important modifications of the model with respect to an otherwise relative standard New Keynesian structure. I estimate this parameter together with the other structural parameters of the model.

1.5.2 Data and estimation settings for quantitative model

For the estimation of the model I retrieve quarterly data for the 7 observables used by Smets and Wouters (2007) (output, consumption, investment, employment, interest rates, wages and inflation) and add the change in nonfinancial sector debt from the flow of funds, scaled by output, as an eighth observable. Consistent with the previous sections of this paper, my data treatment captures explicitly the variation in the relative prices between consumption and investment goods: I obtain real variables by deflating with the consumption deflator of nondurables and services. This is a similar treatment to the one in Justiniano, Primiceri, and Tambalotti (2011). Following the same authors, the sample period for the baseline estimation is 1954:Q3 - 2009:Q1. Details on the data used for estimation are provided in Appendix 1.7.1.

I estimate the model with Bayesian methods, combining the likelihood of the model with prior information on the parameters.[68] I calibrate the means of $\theta_{\pi,t}$ and $\theta_{k,t}$ in the same way as in Section 1.3. For ω I specify a uniform prior between 0 and 1.[69] For comparability to previous studies, I otherwise estimate the same set of parameters as Jermann and Quadrini (2012) and use identical priors. I obtain 1,000,000 draws from a Markov Chain Monte Carlo algorithm and discard the first 20% and use the remaining ones to compute posteriors.

1.5.3 Estimation results: the quantitative role of earnings-based debt

I analyze the role of earnings-based borrowing constraints in the estimated New Keynesian model from a number of angles. Specifically, I present the posterior estimate

[68] An and Schorfheide (2007) provide a survey on Bayesian techniques to estimate DSGE models. For a recent exploration of the sensitivity of these methods to misspecification, see Den Haan and Drechsel (2018).

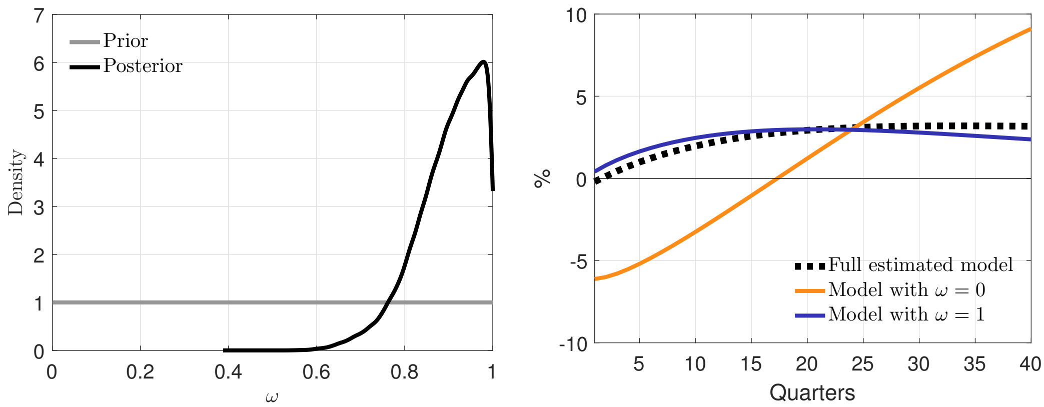
[69] While of course there is prior information available on the shares of earnings based borrowing and debt secured with specific assets – including the evidence shown in the previous parts of this paper – I impose the uniform prior in order to use the model as a separate device to the assign the relative quantitative importance to the constraint components purely based on the information contained in macroeconomic data.

of the weight on the earnings-based component in the constraint, characterize the debt responses to investment and other shocks in different model counterfactuals, and analyze to what degree the different shocks in the model contribute to the variation in the data. In the subsection that follows I turn to studying policy.

The estimated weight on earnings-based debt. Panel (a) of Figure 1.9 plots the prior and posterior density of ω. A value of 0 implies a model with only a collateral constraint, while 1 implies the presence of only an earnings-based constraint. The figure shows that while the prior assigns an equal importance to any weight, the posterior density implies a clear tilt towards the earnings constraint with a mean estimate of $\omega = 0.90$. This finding provides additional evidence that the dynamics in US data, now interpreted through the lens of a richer model structure, favor the earnings-based constraint. The results also highlight that the collateral component remains a feature of the model, although with a much lower weight.[70]

Figure 1.9: PROPERTIES OF THE ESTIMATED QUANTITATIVE MODEL

(a) Weight on earnings-based component (b) Debt IRFs to permanent investment shock



Note: Panel (a) presents the prior and posterior density (grey and black solid lines, respectively) over values of ω, as estimated in the quantitative New Keynesian model on US data. An estimate of 0 implies a model with only a collateral constraint, while an estimate of 1 implies a model with only an earnings-based constraint. See equation (1.18). Panel (b) shows the IRFs to a permanent investment shock, calculated at the posterior means of the estimated model (dotted black line) and for counterfactual models in which the weight of the earnings-based constraint is set to 1 (dark blue line) and 0 (light orange line), but all other parameters are kept at their estimated values. Debt refers to the level of real debt liabilities.

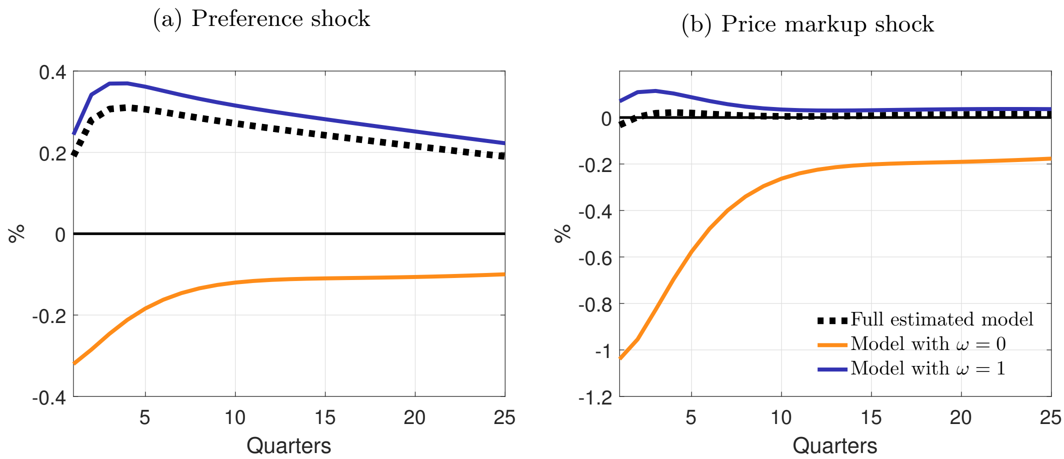
Panel (b) presents the IRFs of real debt liabilities to a permanent positive investment shock, calculated at the posterior means of the estimated model (dotted black line). The chart also contains corresponding IRFs in counterfactual models in which the weight of the earnings-based constraint is set to 1 (dark blue line) and 0 (light orange line), while the other parameters are kept at their posterior mean estimates.[71]

[70]Table 1.13 in the Appendix presents the priors and posterior estimates of all estimated parameters.

[71]Full re-estimation of the other parameters does not change the chart qualitatively, so I stick to presenting this simpler counterfactual experiment. In Section 1.5.4, when focusing on policy shocks, I fully re-estimate all model parameters across the counterfactuals.

In line with the insights of Section 1.3, permanent investment shocks lead to a persistent increase in debt, fostered by a rise in earnings. A pure collateral constrained model would predict a fall in debt, due to the lower value of collateral in equilibrium. The mechanism that is at the heart of Sections 1.3 and 1.4 thus remains intact also when a variety of other frictions are present alongside the borrowing constraint.

Figure 1.10: DEBT IRF TO ADDITIONAL SHOCKS ACROSS MODEL COUNTERFACTUALS



Note: The figure shows the IRFs of firm real debt liabilities to a preference shock (Panel a) and a price markup shock (Panel b). In both cases, the IRFs are calculated at the posterior means of the estimated model (dotted black line) and for counterfactual models which are the weight of the earnings-based constraint is set to 1 (dark blue line) and 0 (light orange line), but all other parameters are kept at their estimated values.

Additional sign differences in debt responses. Figure 1.10 plots the IRFs of firm debt to other selected shocks in the New Keynesian model. Panel (a) presents the responses to a preference shock (an exogenous increase in the household's marginal utility), while Panel (b) shows those for a price markup shock. Again, both charts display the IRF calculated at the posterior means of the estimated model (thick black line) together with corresponding IRFs in counterfactual models in which $\omega = 1$ and $\omega = 0$ (dark blue and light orange lines, respectively). The figure shows that the two alternative borrowing constraints imply opposite signs of the responses of debt also for additional shocks in the model. This is because, similar to the investment shock, both shocks raise earnings but suppress the value of capital. The intuition for the preference shock is that the relative marginal utility between today and tomorrow is raised by the shock. Firms, acting on behalf of the household, cut back on investment, shift resources to the present and pay out dividends. Earnings rise and the capital stock next period is reduced. The intuition for the markup shock is that it allows firms to cut back on production inputs, reducing capital, but simultaneously realize higher profits due to the higher markup. In both cases, the response of the pure earnings-based constraint model lie close to the model in which the weighting between the two components is estimated. This is intuitive, since the posterior estimate of ω is close to, but not equal to, 1.

Variance decomposition of observables. Table 1.3 presents the forecast error variance decomposition of the variables that are used as observables to estimate the model (Table 1.15 in the Appendix shows the corresponding decomposition for a model without any borrowing constraints). This decomposition shows the relative importance that the model attributes to different structural shocks in driving a given observable. For example, according to the model markup shocks to prices and wages are an important driver of inflation dynamics, while consumption growth dynamics are importantly affected by shocks to intertemporal preferences. One observation that stands out in the table is the overall importance of investment shocks. Consistent with similar findings in the literature, the investment margin appears to be crucial for capturing variation in US macroeconomic data (see in particular Justiniano, Primiceri, and Tambalotti, 2010, 2011). This lends further support to my approach of using this shock in the context of studying which type of corporate borrowing constraints are in line with credit dynamics at the macro and micro level.

Table 1.3: VARIANCE DECOMPOSITION OF OBSERVABLES IN QUANTITATIVE MODEL (%)

	TFP	**Inv**	**Pref**	**Price**	**Wage**	**Gov**	**Mon**	**Fin**
Output growth	4.74	53.47	11.7	5.86	2.49	13.09	6.16	2.48
Consumption growth	5.53	5.02	82.81	1.39	1.21	0.02	4.01	0
Investment growth	2.52	86.81	0.25	2.69	2.61	0	5.09	0.03
Inflation	13.07	13.87	4.97	43.48	18.73	0.83	4.98	0.05
Interest rate	4.11	11.94	3.07	16.47	8.12	0.56	55.72	0.01
Employment growth	29.64	39.72	7.27	1.54	3.73	11.12	5.92	1.06
Wage growth	14.21	2.45	2.02	23.86	57.33	0.02	0.08	0.03
Debt issuance	1.13	4.75	0.74	1.65	0.56	0.69	1.14	89.35

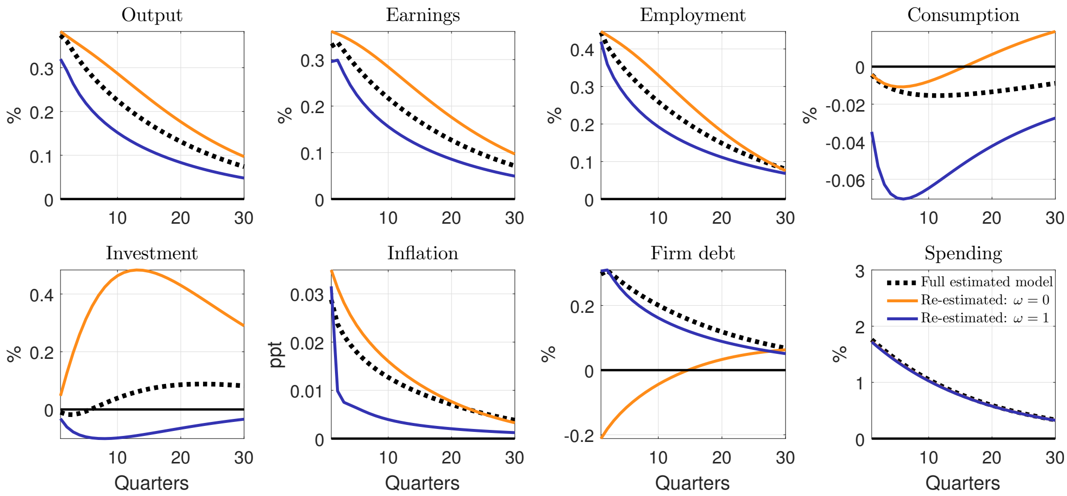
Note: Infinite horizon forecast error variance decomposition of the observables used for the estimation of the model. The decompositions are calculated at the estimated posterior means. Each row presents the decomposition for a given observable, columns correspond to different structural shocks that feature in the model: TFP-Total productivity shock; Inv-Investment shock; Pref-Preference shock; Price-Price markup shock; Wage-Wage markup shock; Gov-Government spending shock; Mon-Monetary policy shock; Fin-Financial shock. Appendix 1.7.7 contains details on the model and specification of the structural shocks.

1.5.4 Counterfactual dynamics for policy shocks

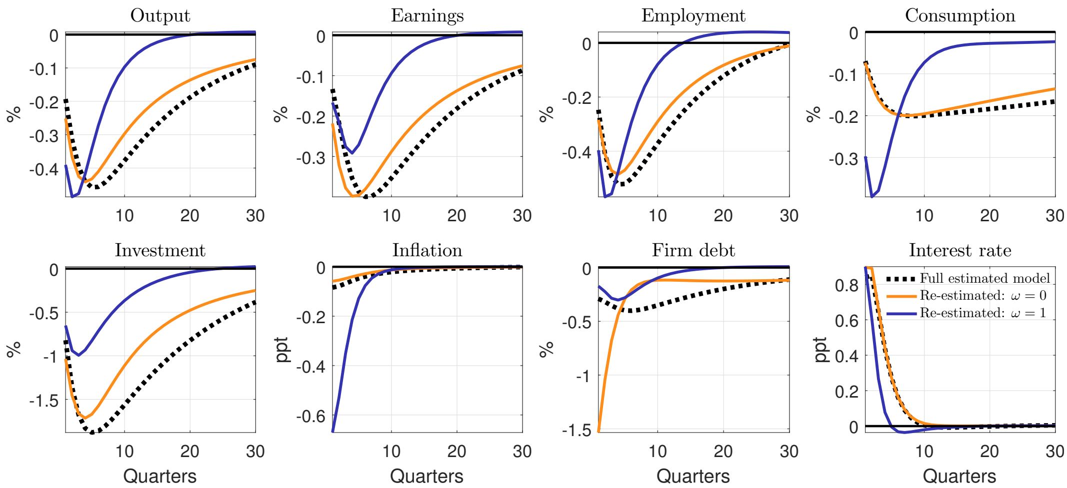
So far my analysis has primarily focused on the dynamics of firm debt to distinguish the role of different borrowing constraints. Figure 1.11 turns to studying the responses of other macroeconomic variables and examines the consequences of the earnings-based constraint for the overall transmission of shocks in the economy. To illustrate the role of the constraints in the quantitative model, I focus on policy shocks. Panel (a) shows the responses to an expansionary fiscal shock, that is, an exogenous increase in government spending. Panel (b) focuses on a contractionary monetary shock, an exogenous increase in the interest rate. In each case I plot the IRFs of the estimated model, calculated at the posterior means as the thick black

Figure 1.11: POLICY SHOCKS IN COUNTERFACTUAL ESTIMATED MODELS

(a) Selected IRFs to an expansionary government spending shock across models



(b) Selected IRFs to a contractionary monetary policy shock across models



Note: The figure shows the IRFs of selected economic variables to a fiscal policy shock (Panel a) and a monetary policy shock (Panel b). In both panels, the IRFs are calculated at the posterior means of the estimated model (dotted black line) and for counterfactual models in which the weight of the earnings-based constraint is set to 1 (dark blue line) and 0 (light orange line), and the other structural parameters are re-estimated. In panel (a), the shock size and persistence to create the IRFs is the same across models. In Panel (b) the size of the shock is adjusted to give the same interest rate response on impact.

line, together with IRFs from counterfactual models based on setting $\omega = 1$ (dark blue line) and $\omega = 0$ (light orange line), and re-estimating the other parameters of the model.[72] For studying policy, this is my preferred type of counterfactual, as I want to characterize hypothetical situations in which a policy maker would only have one or the other model at her disposal.

The figure demonstrates that a policy maker would reach different conclusions across estimated models with alternative borrowing constraints. The presence of earnings-based debt alters the transmission of both fiscal and monetary shocks. In the case of fiscal policy, the intuition is as follows. The spending shock gives rise to temporarily higher demand for the consumption good produced by the firm, which raises earnings and gives an incentive for bringing resources to the presence. Under the earnings-based constraint this incentive is strong enough to crowd out investment. As debt access is determined by earnings, firms borrow more. With a collateral constraint, investment has the additional benefit of building collateral, so firms respond by raising investment and the crowding-out effect disappears. This response is not strong enough to offset the tightening constraint from the fall in the value of capital and the firm reduces its debt position. In net terms, the crowding out of investment dampens the overall stimulus from the spending shock with an earnings constraint. Despite the reduced debt space with the collateral constraint, the stimulus is stronger. Most of the IRFs from the model with both components lie between the other two.

In the case of the monetary shock, the counterfactual estimations show that the earnings-based constraint implies a stronger inflation response and a somewhat stronger but less persistent output response than the collateral constraint. While there are no differences in the sign of any of the responses across models, a comparison of the estimated parameter values of the different models reveals that this is driven by the fact that a model with only an earnings-based constraint implies a relatively low degree of price rigidity relative to the pure collateral model.[73] This finding demonstrates that the specification of the borrowing constraint interacts with other frictions of the model, which highlights that the specification of firm borrowing constraints is crucial for drawing conclusions from a quantitative macro model. Finally, in the case of the monetary shock it is noteworthy that the IRFs of the baseline model, which features both an earnings and collateral component, are generally closer to the ones stemming from the pure collateral constraint model. This is due to the fact that in the re-estimation of the counterfactual models both the dynamics in response to shocks as well as their relative contribution change. This means that any individual IRF of the model with both components does not necessarily lie in

[72] When calculating the IRFs to the fiscal shock, I use the persistence and standard deviations of the disturbances that are implied by the baseline model to ensure comparability. In the case of the monetary shock, I rescale the size of the shock to give the same interest rate response.

[73] Table 1.13 in the appendix shows the priors and posteriors of the model with both components. The model is estimated with a high degree of price adjustment costs, which is driven by the collateral component. The pure collateral constraint is also associated with a high degree of dividend adjustment costs, while, interestingly, estimated investment adjustment costs are similar across the counterfactuals.

between the two counterfactual models. In fact, the presence of the earnings-based constraint makes the responses to monetary shocks stronger, but the implied model assigns a lower importance to monetary policy shocks for US business cycles.

1.6 Conclusion

Capturing the relation between credit and economic activity is crucial for understanding macroeconomic fluctuations. This paper emphasizes the fact that firms' borrowing capacity is tightly connected to their earnings flows, as current earnings are subject to scrutiny by lenders. Grounded on microeconomic evidence on this link, I propose a debt limit that restricts borrowing to a multiple of earnings. The predictions of a business cycle model which features this earnings-based borrowing constraint are in line with both aggregate and cross-sectional credit dynamics in US data. Furthermore, the constraint plays a key role in drawing quantitative conclusions about the transmission of shocks in the economy.

To the extent that debt-to-earnings ratios are targeted in macroprudential regulation, the insights provided in this paper encourage further research to improve policies targeted at firms in credit slumps. Moreover, obtaining a deeper understanding of the cross-sectional heterogeneity that determines the specific conditions under which companies borrow, as well as the potential interaction between different types of credit constraints faced by firms over the business cycle are promising subjects for future research in the field of macro-finance.

1.7 Appendices

1.7.1 Details on the data

This appendix provides details on the data sources used across all sections of the paper. First, Section 1.7.1 describes the Thomson Reuters LPC Dealscan data base and presents summary statistics. This data set is used for the motivational evidence in Section 1.2 of the main text, as well as some of the model calibrations in Section 1.3. Second, the merged data set consisting of the Dealscan data, together with quarterly balance sheet information from Compustat is explained in Section 1.7.1. This data is used in Section 1.4.3 of the main paper, for the local projections of the investment shock in panel data. Third, the construction of the time series data used for the estimation of the SVAR in Section 1.4.1 and the estimation of the quantitative model in Section 1.5 is laid out in Section 1.7.1.

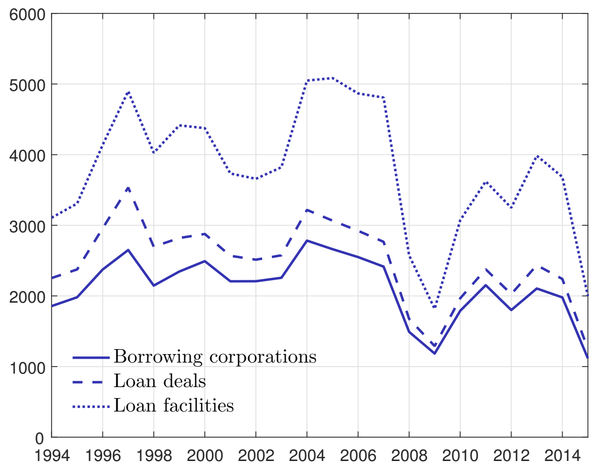
Thomson Reuters LPC Dealscan data set

LPC Dealscan is a detailed loan-level data base provided by Thomson Reuters. The data was retrieved in March 2017 through the LSE Library Services and consists of a full cut of the entire data base provided by Thomson Reuters as of October 2015. The data covers around 75% of the total US commercial loan market (see Chava and Roberts, 2008). The unit of observation is a loan *deal*, sometimes called loan *package*, which can consist of several loan *facilities*. As explained in the main text, rich information is provided both and the deal and facility level. Note that the information is collected at the time of origination but is then not followed over time, so that the data can be thought of as a large cross section with different origination dates.

Data coverage. The raw data set retrieved contains 214,203 deals with 307,660 facilities for 78,646 unique borrowers globally. For the main sample considered in the text I choose loan packages in which the lender is a US nonfinancial Corporation (excluding SIC codes 6000-6999) and the debt is US Dollar denominated. Following Chava and Roberts (2008), I start the sample with loans originated in 1994. These choices result in a sample of 54,400 packages, 83,290 facilities and 15,358 unique borrowing corporations. The number of deals per borrower ranges from 1 to 41, with on average 7.35 deals per borrower. Figure 1.12 summarizes the number of deals, facilities and borrowers split up by origination time.

Summary statistics. Tables 1.4, 1.5, 1.6 and 1.7 provide further descriptive information on the data for the sample described above. Table 1.4 provides summary statistics on the size of both deals and facilities and of the maturity of the loans, which is available at the facility level. As the table shows loans reach from single digit million amounts up to the size of a few billion dollars. Facility amounts are smaller on average, which is true by construction since a deal consists of at least one facility. The maturity of a facility is on average between 4 and 5 years (52 months). 1.5 shows the coverage of the data across industries. Table 1.6 lists the ten most frequently stated loan purpose, which is provided at the deal level. This information is available for every deal in the sample (no missing fields), although it is apparent that the number one category "corporate purpose" is relatively unspecific. Table 1.7 lists the most common asset *types* of collateral pledged in secured loan facilities.

Figure 1.12: COVERAGE OF DEALSCAN SAMPLE BY ORIGINATION DATE



Note: The figure plots the number of loan deals (or packages), loan facilities and borrowing corporations for the sample used in the main analysis of the paper, broken down by origination date since 1994. The sample covers USD denominated debt for US nonfinancial corporations.

Table 1.4: SUMMARY STATISTICS FOR DEALSCAN DATA

	Deal amount (mio 2009 USD)	**Facility amount (mio 2009 USD)**	**Facility maturity (months)**	**Interest rate (drawn spread)**
Mean	418.2	273.2	52	259
Std. deviation	1002.1	683.1	27	166
1st percentile	2.5	1.3	5	20
10th percentile	23.7	10.4	12	65
25th percentile	60.0	29.9	36	150
Median	151.2	92.2	60	250
75th percentile	395.8	257.4	60	330
90th percentile	951.1	619.4	84	450
99th percentile	4144.2	2750.0	120	830
Observations	54,397	83,288	76,205	70,282

Note: Summary statistics for Dealscan loan sample used for the main analysis in the paper. Real values were obtained using the US business deflator with base year 2009. The interest rate in the all-in spread for drawn facilities, expressed as a spread over LIBOR in basis points. Changes in the number of observation result from missing fields.

Table 1.5: INDUSTRY COVERAGE IN DEALSCAN DATA

Industry	**No of firms**	**No of loan deals**	**Amount borrowed**
Consumer Nondurables	1,120	4,420	1.83
Consumer Durables	424	1,738	0.80
Manufacturing	1,741	7,036	2.52
Oil, Gas, and Coal	805	3,479	1.78
Chemicals	382	1,699	0.91
Business Equipment	1,503	4,718	1.76
Telephone and TV	795	2,755	2.21
Utilities	767	3,964	2.27
Wholesale, Retail	2,216	8,579	2.83
Healthcare	1,003	3,469	1.65
Other	3,311	10,982	3.93
No SIC code available	1,290	1,560	0.25

Note: Industries are based on the Fama-French 12 Industry Classification. Finance and Utilities have been excluded. The amount borrowed is in trillions of 2009 real USD.

Table 1.6: FREQUENCY OF STATED DEAL PURPOSE IN DEALSCAN DATA

Deal purpose	**Share (equal-weighted)**	**Share (value-weighted)**
Corporate purposes	46.7%	44.0%
Working capital	12.3%	7.6%
Debt Repayment	11.9%	9.6%
Takeover	6.3%	13.8%
Acquisition line	5.3%	4.2%
LBO	4.4%	4.9%
CP backup	3.8%	8.1%
Dividend Recap	1.4%	1.1%
Real estate	1.3%	0.3%
Debtor-in-possession	1.0%	0.5%

Note: The table shows the ten most frequently stated "deal purposes". This information is available at the deal level for all 50,437 observations in the US sample. The first column calculates the frequency by firm, the second one by (real) USD.

Table 1.7: MOST FREQUENTLY PLEDGED ASSETS IN SECURED LOANS IN DEALSCAN DATA

Collateral type	**Number of loan facilities**	**Volume in bn USD**
Property & Equipment	2292	353
Accounts Receivable and Inventory	1801	332
Intangibles	1367	238
Cash and Marketable Securities	989	328
Real Estate	737	142
Ownership of Options/Warrants	104	19
Patents	84	12
Plant	50	12
Agency Guarantee	25	6

Note: The numbers in this table are calculated by restricting Dealscan facilities to secured facilities and then calculating the frequencies of different security types. The table focuses on *specific* asset categories, i.e. excludes the categories "unknown", "all", and "other". According to Lian and Ma (2018), facilities secured by all assets (excluded in this table), can generally be classified as cash-flow based loans, as the value of this form of collateral in the event of bankruptcy is calculated based on the cash flow value from continuing operations. The key function of having security is to establish priority in bankruptcy.

Merged Dealscan-Compustata panel data set

Compustat Northamerica Quarterly. This data set provides accounting data for publicly held companies in the US and Canada at quarterly frequency starting in 1960. The data was accessed through the Upenn Whartson Research Data Services (WRDS) in September 2016. I keep firms incorporated in the United States with positive assets and sales and exclude Financials (SIC codes 6000-6999). In addition, I generally exclude the sector of 'unclassifiable' firms (SIC codes starting with 99), since this sector contains very few large holding firms, which are typically financial firms (e.g. Berkshire Hathaway). Finally I drop firms that are present less than 5 years. These sample restrictions are typically made in papers that focus on nonfinancial Compustat firms (see for example Bates et al., 2009).

Merge of Dealscan with Compustat. As described in the text, I use Michael Roberts' identifier link, which is available on Michael Roberts' personal website and which is infrequently updated. See also Chava and Roberts (2008). The version of the link file which I retrieved is the April 2018 version. I drop firms from Compustat that do not appear at least once in the Dealscan data and restrict the sample to the period covered by the link file. I deseasonalise the variables I use from Compustat by regressing them on quarter-dummies before using them in the actual regressions. The resulting merged data set covers more than 150,000 firm-quarter observations for more than 4,000 distinct firms from 1994 to 2015.

Summary statistics for the merged data set. Table 1.8 provides summary statistics for the firms in the full Compustat-Dealscan panel, which is constructed as described above, and used to estimate equation (1.16). Table 1.9 presents the corresponding information for firms based on the baseline classification used in equation (1.17). Note that since firms can have several loan issuances, a given firm may appear in several panels of the table. For *a given time period* in the estimation of (1.17), the grouping is mutually exclusive.

Table 1.8: SUMMARY STATISTICS FOR FULL COMPUSTAT-DEALSCAN PANEL

	Firm-qrt obs	**Mean**	**SD**	**Min**	**Median**	**Max**
Real total assets (bn 2009 USD)	153,554	4.6	16.2	0.0	0.8	542.7
Real sales (bn 2009 USD)	153,554	1.0	3.7	0.0	0.2	124.3
Real sales growth (percent)	149,049	3.4	16.6	-27.6	1.9	43.3
Employment (thousands)	136,575	14.3	53.5	0.0	2.8	2200.0
Real debt liabilities (bn 2009 USD)	153,554	1.4	6.4	0.0	0.2	339.6
Cash ratio	153,543	0.1	0.1	0.0	0.0	0.9
Market-to-book ratio	140,325	1.8	1.8	0.5	1.4	45.0
Book leverage (broad)	153,543	0.6	0.2	0.1	0.6	1.3
Book leverage (narrow)	153,543	0.4	0.2	0.0	0.3	0.9

Table 1.9: SUMMARY STATISTICS FOR SUBGROUPS IN COMPUSTAT-DEALSCAN PANEL

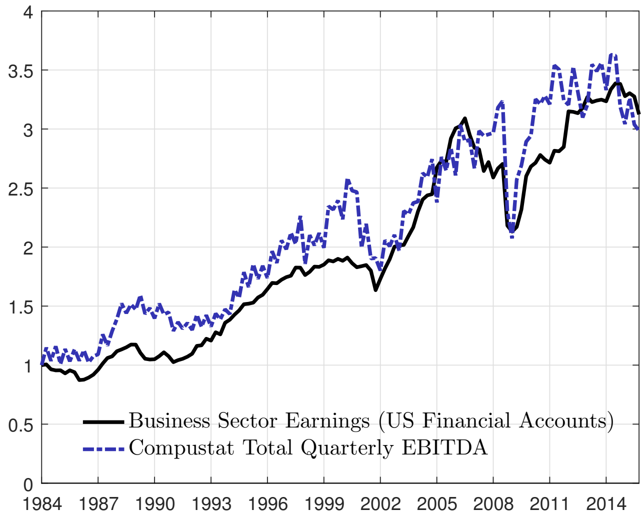
	Firm-qrt obs	**Mean**	**SD**	**Min**	**Median**	**Max**
Panel (a): Borrowers taking at least one loan with earnings covenants only ($N = 1,721$)						
Real total assets (bn 2009 USD)	46,680	5.4	17.2	0.0	1.6	455.6
Real sales (bn 2009 USD)	46,680	1.1	2.7	0.0	0.4	55.0
Real sales growth (percent)	46,044	4.9	16.3	-27.6	2.8	43.3
Employment (thousands)	43,164	17.7	40.8	0.0	5.4	707.9
Real debt liabilities (bn 2009 USD)	46,680	1.8	6.1	0.0	0.4	251.9
Cash ratio	46,668	0.1	0.1	0.0	0.0	0.9
Market-to-book ratio	43,848	1.7	1.0	0.5	1.5	16.8
Book leverage (broad)	46,668	0.6	0.2	0.1	0.6	1.3
Book leverage (narrow)	46,668	0.4	0.2	0.0	0.3	0.9
Panel (b): Borrowers taking at least one loan with specific collateral only ($N = 1,470$)						
Real total assets (bn 2009 USD)	28,128	3.5	10.2	0.0	0.6	192.8
Real sales (bn 2009 USD)	28,128	0.8	3.0	0.0	0.1	86.3
Real sales growth (percent)	26,652	4.7	17.6	-27.6	2.8	43.3
Employment (thousands)	25,860	12.5	52.6	0.0	2.1	1900.0
Real debt liabilities (bn 2009 USD)	28,128	1.5	4.4	0.0	0.2	131.1
Cash ratio	28,128	0.1	0.1	0.0	0.0	0.9
Market-to-book ratio	25,428	1.7	1.5	0.5	1.3	45.0
Book leverage (broad)	28,128	0.7	0.3	0.1	0.7	1.3
Book leverage (narrow)	28,128	0.5	0.3	0.0	0.4	0.9
Penal (c): Borrowers taking at least one loan with both ($N = 1,855$)						
Real total assets (bn 2009 USD)	44,124	2.2	9.8	0.0	0.6	513.3
Real sales (bn 2009 USD)	44,124	0.5	1.3	0.0	0.1	51.9
Real sales growth (percent)	42,864	6.0	17.8	-27.6	3.5	43.3
Employment (thousands)	41,652	9.2	24.0	0.0	2.6	355.0
Real debt liabilities (bn 2009 USD)	44,124	1.0	5.6	0.0	0.2	307.5
Cash ratio	44,124	0.1	0.1	0.0	0.0	0.9
Market-to-book ratio	40,764	1.6	0.9	0.5	1.3	12.0
Book leverage (broad)	44,124	0.6	0.2	0.1	0.6	1.3
Book leverage (narrow)	44,124	0.5	0.3	0.0	0.5	0.9
Panel (d): Borrowers taking at least one loan without either ($N = 844$)						
Real total assets (bn 2009 USD)	20,424	12.8	26.4	0.0	4.2	375.8
Real sales (bn 2009 USD)	20,424	2.6	5.6	0.0	0.7	66.0
Real sales growth (percent)	20,040	4.7	17.8	-27.6	2.7	43.3
Employment (thousands)	14,724	39.4	83.9	0.0	10.3	1383.0
Real debt liabilities (bn 2009 USD)	20,424	3.8	10.2	0.0	1.2	216.3
Cash ratio	20,424	0.1	0.1	0.0	0.0	0.9
Market-to-book ratio	18,048	1.7	1.0	0.5	1.4	12.7
Book leverage (broad)	20,424	0.6	0.2	0.1	0.6	1.3
Book leverage (narrow)	20,424	0.4	0.2	0.0	0.3	0.9

US aggregate time series data

Data sources. The aggregate time series data used for the SVAR analysis and the estimation of the quantitative model come from a number of sources, including the Bureau of Economic Analysis, the Bureau of Labor Statistics and the US Financial Accounts provided by the Federal Reserve (also known as Flow of Funds). I retrieved these series using FRED and the data download program of the US Financial Accounts. In the treatment of relative prices in both panels, I closely follow Fisher (2006) and Justiniano, Primiceri, and Tambalotti (2011). The selection of variables for the New Keynesian model is the same as Jermann and Quadrini (2012). Table 1.10 lists the time series and their construction, together with the specific identifiers.

Details on the earnings measure. To calculate an aggregate corporate earnings/profit measure, I use the item 'FA146110005.Q: Income before taxes' for the nonfinancial business sector, available from the table F.102 in the US Financial Accounts. I cross-checked the cyclical properties of this series with the 'ebitda' item from Compustat and found it to be relatively similar, see Figure 1.13 below:

Figure 1.13: US FINANCIAL ACCOUNTS VS COMPUSTAT



Note: The figure shows a comparison of earnings measures from the US financial accounts and Compustat Quarterly. Both series are normalized to 1 in 1984:Q1. The Compustat series is not seasonally adjusted.

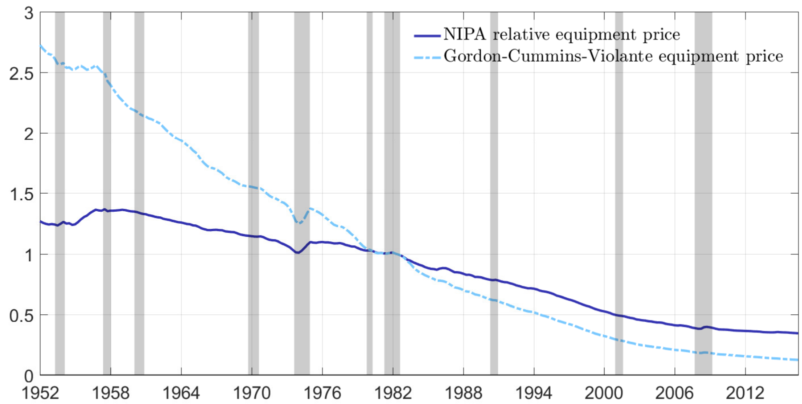
Table 1.10: Details on aggregate time series data

Variable	Series sources and construction	Transform
	Panel (a): Data used in estimation of SVAR	
Variable	**Series sources and construction**	**Transform**
Relative price of investment	Implicit price deflator of nonresidential fixed equipment investment (FRED: Y033RD3Q086SBEA), deflated with implicit price deflator of personal consumption expenditures of nondurable goods and services (FRED: CONSDEF)	log diff
Relative price of investment (alternative measure)	See DiCecio (2009) for details (FRED: PERIC)	log diff
Labor productivity	Nominal business sector value added (FRED: A195RC1Q027SBEA), deflated with consumption deflator (see above), divided by hours worked (see below)	logdiff
Hours worked	Hours of all persons in the nonfarm business sector (FRED: HOANBS)	log
Business sector earnings	Sum of nominal income before taxes in the nonfinancial noncorporate sector (USFA: FA146110005.Q) and corporate profits before tax excluding IVA and CCAdj (USFA: FA146110005.Q), deflated with consumption deflator (see above)	logdiff
Level of the capital stock	Constructed from capital expenditures in the nonfinancial business sector (USFA: FA145050005.Q) minus depreciation (consumption of fixed capital in the nonfinancial business sector, USFA: FA106300083.Q), valued at the relative price of investment (see above)	logdiff
Business sector debt	Level of debt securities and loans in the nonfinancial bussiness sector (constructed from USFA: FA104122005.Q and FA144123005.Q), deflated with consumption deflator (see above)	logdiff
	Panel (b): Data used in estimation of New Keynesian model	
Variable	**Series sources and construction**	**Transform**
Output	Nominal GDP (FRED: GDP), divided by population (FRED: B230RC0Q173SBEA), deflated with consumption deflator (see above)	logdiff
Consumption	Real consumption expenditures of nondurable goods and services (FRED: PCNDGC96 and PCESVC96), divided by population (see above)	logdiff
Investment	Sum of nominal gross private domestic investment expditures (FRED: GPDI) and nominal private consumption expenditures on durable goods (FRED: PCDG), divided by population (see above), deflated with consumption deflator (see above)	logdiff
Hours worked	See above	logdiff
Real wage	Nominal compensation per hour in the nonform business sector (FRED: COMPNFB), deflated with consumption deflator (see above)	logdiff
Inflation	Percentage change in consumption deflator (see above)	none
Interest rate	Nominal effective Federal Funds Rate (FRED: FEDFUNDS)	none
Debt issuance / output	Change in level of business sector debt (sum of USFA: FA104122005.Q and FA144123005.Q), divided by real output (see above)	none

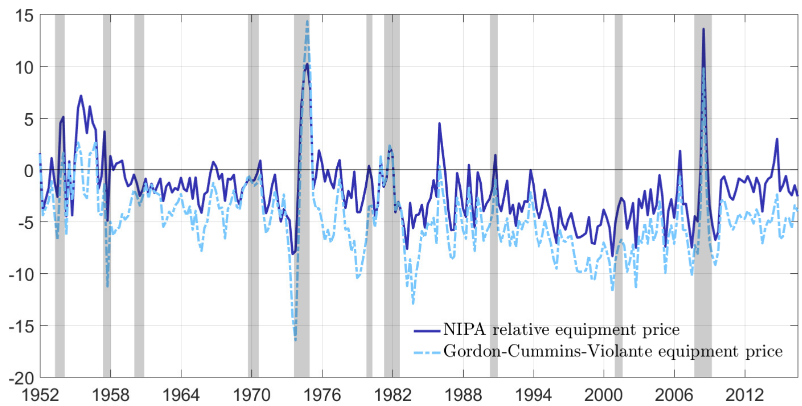
Details on relative equipment prices. Figure 1.14 compares the two alternative measures used for the relative price of equipment investment. The first is the one based on NIPA data, constructed as the ratio between the equipment investment deflator and the deflator of consumption on nondurables and services. The second one is the Gordon-Violante-Cummins (GVC) relative equipment price, see Cummins and Violante (2002) and DiCecio (2009). Panel (a) plots the evolution in the level and Panel (b) plots the quarterly growth rates. More details can be found in Table 1.10.

Figure 1.14: MEASURES OF THE RELATIVE EQUIPMENT PRICE

(a) Levels (1982:Q3 = 100)



(b) Growth rates (annualized %)



Note: Panel (a) plots the evolution in the level and Panel (b) the quarterly growth rates of the two alternative measures used for the relative price of equipment. The solid dark blue line shows the one constructed from NIPA deflators and the dashed light blue one the Gordon-Violante-Cummins (GVC) relative equipment price, see Cummins and Violante (2002) and DiCecio (2009). Table 1.10 contains additional details.

Table 1.11 reports the results from an augmented Dicker-Fuller (ADF) test on the two alternative equipment price series plotted in Figure 1.14. The test is specified as in Gali (1996). The model under the null has a unit root, the alternative is the same model with drift and deterministic trend. The lag order is 4. Consistent with the assumptions required by the SVAR identification scheme, the test fails to reject a unit root in the level, but rejects a unit root in after first-differencing for both alternative measures.

Table 1.11: RESULTS OF UNIT ROOT TESTS ON EQUIPMENT PRICE SERIES

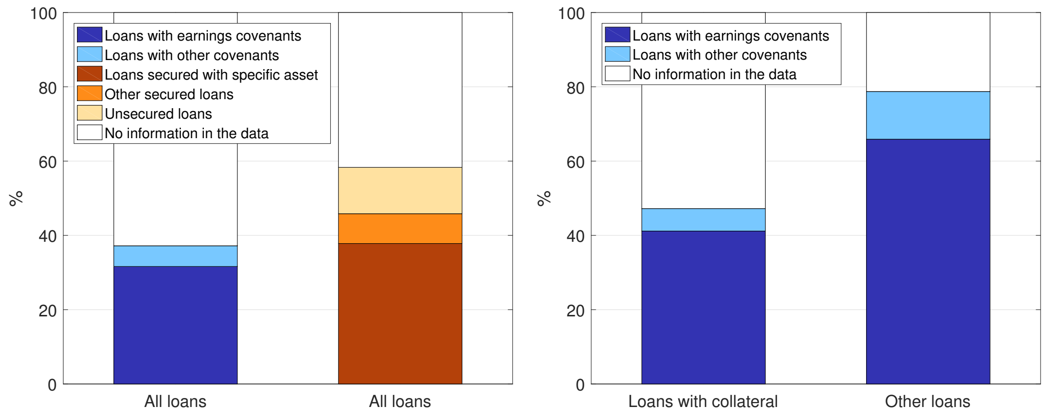
	Test statistic	**5% critical value**	**Reject?**
NIPA levels	-3.34	-3.43	No
NIPA first differences	-5.40	-3.43	Yes
GVC levels	-0.15	-3.43	No
GVC first differences	-6.99	-3.43	Yes

Note: Unit root test on alternative equipment price series in levels and first differences. See Table 1.10 for details on the series. Following Gali (1996) the table reports the relevant t-statistics for the null hypothesis of a unit root in the level and the first difference of each time series, based on an augmented Dicker-Fuller (ADF) test with 4 lags, intercept and time trend.

1.7.2 Additional evidence

Figure 1.15: THE IMPORTANCE OF EARNINGS-BASED AND ASSET-BASED DEBT (EQUAL-WEIGHTED SHARES)

(a) Frequencies of covenants and secured loans (b) Covenants within (un)collateralized loans



Note: The figure repeats Figure 1.1 of the main text for equal-weighted rather than value-weighted shares. Panel (a) displays the shares of loan deals that contain covenants (left bar) and are secured/unsecured (right bar). In the left bar, the dark blue area represents the share with at least one earnings-based covenant. The light blue area covers loans with covenants unrelated to earnings. In the right bar, the different orange shades capture loans secured with specific assets (dark), other secured loans (medium) and unsecured loans (light). In both bars, loans without the relevant information are represented by the white area. Panel (b) repeats the left column of Panel (a), but breaks down the sample into loans secured with specific assets and other loans (with any information on secured/unsecured). The sample used for both panels consists of loan deals issued between 1994 and 2015 by US nonfinancial corporations.

1.7.3 Discussion of microfoundation

The two borrowing constraints introduced in Section 1.3 of the text are exogenously imposed on the firm. This appendix discusses a formal rationalization of these constraints. I lay out a setting in which the constraints are derived as the solution to an enforcement limitation, in which borrower and lender predict the renegociation outcomes in the event of a default. The appendix also provides a further discussion of the potential frictions underlying the earnings-based constraint, by giving a summary of the literature on the microfoundations of loan covenants and presenting additional details on regulatory requirement in relation to earnings covenants.

A formal rationalization of the alternative borrowing constraints

Collateral constraint. I begin with this constraint, as it is more familiar in the literature. Consider the firm as described in the text and the first type of debt it has access to. Suppose that at the end of period t, when all transactions have been settled, the firm can default on its debt liabilities, which at this point amount to $\frac{b_{k,t}}{1+r_{k,t}}$. In the absence of any punishment, the firm would have an advantage from doing this, as the repayment of $b_{k,t}$ would not reduce resources in its flow of dividends constraint (1.4) next period.

Suppose the legal environment surrounding this type of debt is such that in the event of default the lender can address a court which grants it the right to seize the firm's collateral at the beginning of $t+1$. The lender will be able to re-sell this collateral after depreciation at market prices, but incur a transaction cost which is a fraction $(1-\theta_k)$ of the resale value of capital. Hence, instead of having $\frac{b_{k,t}}{1+r_{k,t}}$ on the asset side of her balance sheet at the end of the period, the lender now has a legal claim on selling the asset tomorrow, which is valued as $\theta_k \mathbb{E}_t p_{k,t+1}(1-\delta)k_t$. If the collateral is seized by the lender, the firm is required to stop operating.

Suppose that before going to the next period, lender and borrower are able to renegotiate. The borrower can offer a settlement payment $s_{k,t}$ to the lender, in combination with a promise to repay the amount of liabilities she has defaulted on. Any settlement amount that the lender would agree to needs to satisfy

$$s_{k,t} + \frac{b_{k,t}}{1+r_{k,t}} \geq \theta_k \mathbb{E}_t p_{k,t+1}(1-\delta)k_t. \tag{1.19}$$

Now, for the firm to never choose to default, the value of operating in absence of default must exceed the value of the firm after successful renegotiation. In other words, as long as the required settlement payment is positive, the predicted outcome of renegotiation is such that the firm would never choose to default. Formally, from combining this non-negativity condition with (1.19), we obtain

$$
\begin{aligned}
s_{k,t} &\geq 0 && (1.20)\\
\theta_k \mathbb{E}_t p_{k,t+1}(1-\delta)k_t - \frac{b_{k,t}}{1+r_{k,t}} &\geq 0, && (1.21)
\end{aligned}
$$

which can be rearranged to equation (1.9) in the text.

Earnings-based constraint. Suppose that for the second debt type the environment is such that when the firm defaults on its liabilities $\frac{b_{\pi,t}}{1+r_{\pi,t}}$ at the end of $t+1$, the court grants the lender the right to seize ownership of the entire firm. She can then either operate the firm herself or sell it on the market. Importantly, however, the lender is uncertain about the value of the firm in this case. Denote $\tilde{V}_{d,t}^{end}$ the end-of-period value of the firm after ownership rights have been transferred to the lender. In order to determine this uncertain value, the lender uses the common practice of valuation by multiples.[74] Specifically, she evaluates firm ownership after default by using fixed multiple of the last available realization of a fundamental profitability indicator, EBITDA. Formally,

$$
\tilde{V}_{d,t}^{end} \approx \theta_\pi \pi_t. \tag{1.22}
$$

In this case, the required settlement amount in the renegotiation process needs to satisfy

$$
\begin{aligned}
s_{\pi,t} &\geq 0 && (1.23)\\
\theta_\pi \pi_t - \frac{b_{k,t}}{1+r_{k,t}} &\geq 0. && (1.24)
\end{aligned}
$$

The last inequality can be arranged to (1.8) in the text.

Remarks. As shown above, both collateral and earnings-based borrowing constraint can arise in a world of limited enforcement. Specifically, they can be derived from a situation in which lenders and borrowers predict the outcome of a renegotiation process that would be triggered in the event of default. Based on the predicted outcomes of this renegotiation, the firm will not choose to default, but borrowing is subject to the respective limit on the debt liabilities.

In the setting laid out, the underlying contractual frictions behind equations (1.8) and (1.9) differ as follows. In the case of the earnings-based constraint, there is an informational friction regarding the contingent firm value. The transfer of ownership rights is not accompanied by a transaction cost, but by uncertainty that surrounds the value of the firm after ownership rights have been transferred. In the case of collateral, there is a rational prediction of the resale value, but a transaction cost needs to be incurred.

[74] For a textbook treatment, see Damodaran (2012).

Further discussion of the flow-based constraint

Microfoundation of loan covenants in the literature. Since I empirically motivated the earnings-based constraint based on the presence of loan covenants, studying the academic literature that has studied these covenants lets us get a sense of how researchers conceptualize earnings-based constraints at a micro level. As I stress in Section 1.2 of the text, however, covenants are one but not the only mechanism through which current earnings flows feed back to the ability to issues debt.

The literature on loan covenants can broadly be distinguished between two strands. The first are empirical papers that investigate covenants and their economic effects in firm-level data. This includes the papers that I have cited in Section 1.2 of the text. Key references are for example Chava and Roberts (2008), Roberts and Sufi (2009a) and Bradley and Roberts (2015). These papers do not provide a fully fledged theoretical rationalization of why loans contain covenants, but mostly take them as a given empirical phenomenon and test their effects in the data. Nevertheless these papers typically do provide some remarks on the rationale for covenants to guide their analysis. The second strand is theoretical work in the (incomplete) contracts literature that directly addresses the microfoundation of covenants. This literature builds on seminal work of Aghion and Bolton (1992) and goes back at least to Jensen and Meckling (1976). One example that directly studies the contractual design of covenants is Garleanu and Zwiebel (2009).

Both streams of work have generally highlighted moral hazard issues. A compact description is provided by Chava and Roberts (2008). According to the authors a key rationale for covenants is the allocation of contingent control rights over the firm. Adding covenants to a contract provide debt holders with the option to intervene in the companies management. In the same spirit, Dichev and Skinner (2002) refer to covenants as "trip wires". Such a contingent transfer of control rights provides an additional incentive to management behavior that is in line with the debt holders' objectives. While in my macro model these moral hazard problems are not explicitly present, the formal rationalization above has shown that is possible to generate the constraint from an enforcement issue. Furthermore, the earnings-based constraint introduces an important feedback between firms' earnings and their ability to borrow. The fact that the covenants literature finds large economic effects of covenants (and their breaches) on the borrowing firm suggests that such a feedback is a plausible empirical pattern.

Regulation. As mentioned in the main text, an alternative way to think about the earnings-based constraint is the presence of regulation that lenders, in particular banks, are subject to. For example, regulators in the US define "leveraged transactions", among other criteria, based on the debt-to-EBITDA ratio of borrowers.[75]

[75] See for example the *US Interagency Guidance on Leveraged Lending (2013)*, available at https://www.federalreserve.gov/supervisionreg/srletters/sr1303a1.pdf. Similar definitions can be found in EU regulations.

Whether transactions are defined in this way in turn affects risk-weights and hedging requirements for lenders.

In the case of mortgages, regulatory requirements on income flows have been highlighted by Greenwald (2017), who also studies collateral (loan-to-value) and flow-related (payment-to-income) constraints. He imposes the two borrowing constraints household debt and refers to them as “institutional rules that are not the outcome of any formal optimization problem”. Given that both collateral and the debt-to-EBITDA ratio also feature in the regulation of lenders that provide fund to nonfinancial firms, an alternative way to think about equations (1.8) and (1.9) is that they are the outcome regulation rather than an underlying contracting frictions that lender and borrowing need to overcome.

1.7.4 Details on the model of Section 1.3

Firm optimality conditions

The firm's optimality conditions with respect to n_t, $b_{k,t}$, $b_{\pi,t}$ and k_t and i_t are derived as follows[76]

$$F_{n,t} = w_t \tag{1.25}$$

$$R_{k,t}\mathbb{E}_t\left\{\frac{\Lambda_{t+1}}{\Lambda_t}\right\} + \mu_{k,t}\frac{R_{k,t}}{1+r_{k,t}} = 1, \tag{1.26}$$

$$R_{\pi,t}\mathbb{E}_t\left\{\frac{\Lambda_{t+1}}{\Lambda_t}\right\} + \mu_{\pi,t}\frac{R_{\pi,t}}{1+r_{\pi,t}} = 1, \tag{1.27}$$

$$Q_t = \mathbb{E}_t\left\{\frac{\Lambda_{t+1}}{\Lambda_t}\left[(1-\delta)Q_{t+1} + F_{k,t+1} + \mu_{\pi,t+1}\theta_\pi F_{k,t+1}\right] + \mu_{k,t}\theta_k(1-\delta)p_{k,t+1}\right\} \tag{1.28}$$

$$Q_t v_t[(1-\Phi_t) - \Phi_{1,t}i_t] + \mathbb{E}_t\left\{\frac{\Lambda_{t+1}}{\Lambda_t}Q_{t+1}v_{t+1}\Phi_{-1,t+1}i_{t+1}\right\} = 1 \tag{1.29}$$

where $F_{n,t}$ and $F_{k,t}$ denote the marginal products of labor and capital, respectively. The Lagrange multipliers on the borrowing constraints (1.8) and (1.9) are denoted by $\mu_{\pi,t}$ and $\mu_{k,t}$, respectively. Q_t is the Lagrange multiplier on the capital accumulation equation (1.3) and defines the market value of the capital stock (see Hayashi, 1982). As is typical in models with adjustment costs, its dynamics are characterized by the first order condition of investment, equation (1.29). In this equation $\Phi_{1,t}$ and $\Phi_{-1,t+1}$ denote the partial derivatives of $\Phi_t\left(\frac{i_t}{i_{t-1}}\right)$ and $\Phi_{t+1}\left(\frac{i_{t+1}}{i_t}\right)$ to i_t, respectively. The capital price $p_{k,t}$ that is relevant in the collateral constraint is given by (1.10) in the main text.

Household, government, and definition of equilibrium

Household problem

The household's objective is to maximize expected discounted lifetime utility

$$\mathbb{E}_0\sum_{t=0}^{\infty}\beta^t u(c_t, n_t), \tag{1.30}$$

subject to the budget constraint

$$c_t + \frac{b_{\pi,t}}{1+r_{\pi,t}} + \frac{b_{k,t}}{1+r_{k,t}} + p_t s_t + T_t = w_t n_t + b_{\pi,t-1} + b_{k,t-1} + s_{t-1}(d_t + p_t). \tag{1.31}$$

Equity shares in the firm are denoted by s_t and evaluated at price p_t. T_t is a lump sum tax. I specify preferences using a log-log utility function in consumption and leisure

$$u(c_t, n_t) = log(c_t) + \chi\, log(1-n_t), \tag{1.32}$$

[76]For ease of notation I focus on the case without dividend adjustment costs ($\psi = 0$).

where χ governs the relative utility of leisure. The household takes $r_{k,t}, r_{\pi,t}, p_t$ and w_t as given when maximizing her objective.

Household optimality conditions. The household's optimality conditions with respect to n_t, $b_{k,t}$, $b_{\pi,t}$ and s_t are

$$u_{c_t} w_t + u_{n_t} = 0 \tag{1.33}$$
$$u_{c_t} = \beta(1 + r_{k,t})\mathbb{E}_t u_{c_{t+1}} \tag{1.34}$$
$$u_{c_t} = \beta(1 + r_{\pi,t})\mathbb{E}_t u_{c_{t+1}} \tag{1.35}$$
$$u_{c_t} p_t = \beta\mathbb{E}_t(d_{t+1} + p_{t+1})u_{c_{t+1}}, \tag{1.36}$$

where u_{c_t} and u_{n_t} denote marginal utility of consumption and labor, respectively.

Government

The lump sum tax T_t is required to finance the tax advantage of debt that is given to the firm, which amounts to the difference between debt issued (valued at $R_{j,t}^{-1}$) and debt received (valued at $(1 + r_{j,t})^{-1}$) for both debt types $j \in \{k, \pi\}$. In principle this lump sum tax could be levied on the firm as well, which would not alter the results. For simplicity I assume that the government does not save or borrow. Taken together, budget balance requires

$$T_t = \frac{b_{k,t}}{R_{k,t}} - \frac{b_{k,t}}{(1 + r_{k,t})} + \frac{b_{\pi,t}}{R_{\pi,t}} - \frac{b_{\pi,t}}{(1 + r_{\pi,t})}. \tag{1.37}$$

Equilibrium

I collect the exogenous states of the model in the vector $\mathbf{x}_t = (z_t, v_t, \phi_t)'$. These variables are assumed to follow a stochastic process of the form $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{u}_t$, which will be specified in the parameterization section below. The endogenous states of the model are k_{t-1}, $b_{k,t-1}$ and $b_{\pi,t-1}$. A dynamic competitive equilibrium is then defined as a set of quantities $\{d_t, n_t, b_{k,t}, b_{\pi,t}, k_t, c_t, s_t, T_t\}_{t=0}^{\infty}$ and prices $\{w_t, Q_t, p_{k,t}, R_{k,t}, R_{\pi,t}, r_{k,t}, r_{\pi,t}, \mu_{k,t}, \mu_{\pi,t}, \Lambda_t\}_{t=0}^{\infty}$ such that:

1. $d_t, n_t, b_{k,t}, b_{\pi,t}$ and k_t solve the firm's maximization problem specified above
2. $c_t, n_t, b_{k,t}, b_{\pi,t}$ and s_t solve the household's maximization problem specified above
3. The household owns the firm: $\Lambda_t = \beta^t u_{c_t}$ and $s_t = 1$
4. The government's budget constraint holds
5. The exogenous disturbances follow $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{u}_t$
6. Markets clear

The equilibrium admits a recursive formulation, to which the solution is a set of policy functions that map state variables into endogenous controls. Section 1.7.4 of this appendix contains details on the calculation of the model's steady state. I solve for the policy functions with standard first-order perturbation techniques.

Specification of stochastic processes

The stochastic processes underlying the exogenous disturbances are defined as

$$
\begin{aligned}
log(z_t) &= (1-\rho_z)log(\overline{z}) + \rho_z log(z_{t-1}) + u_{z,t} && (1.38)\\
log(v_t) &= (1-\rho_v)log(\overline{v}) + \rho_v log(v_{t-1}) + u_{v,t} && (1.39)\\
log(\phi_t) &= (1-\rho_\phi)log(\overline{\phi}) + \rho_\phi log(\phi_{t-1}) + u_{\phi,t} && (1.40)
\end{aligned}
$$

where the structural shocks $\{u_{z,t}, u_{v,t}, u_{\phi,t}\}$ are uncorrelated, iid, mean zero, normally distributed random variables with standard deviations $\{\sigma_z, \sigma_v, \sigma_\phi\}$.

Sketch of analytical calculation of the steady state

To compute the steady state of the model, I proceed as follows:

1. Drop time subscripts, obtain a system in steady state variables.
2. Steady state must fulfill $r_j = (1-\beta)/\beta$, $R_j = 1 + r(1-\tau_j)$ and $\mu_j = (1 + r_f)(1/R_j - \beta)$ from bond Euler equations for firm and household, that is, equations (1.26), (1.27), (1.34) and (1.35).
3. Steady state must fulfill $Q = 1$
4. Solve (1.28) for the steady state capital-labor ratio as a function of model primitives.
5. Calculate steady state wage rate w from (1.25) using steady state capital-labor ratio.
6. Combine the capital-labor ratio, the wage rate, (1.33) and the resource constraint to calculate n as a function of model primitives.
7. Recover k from the definition of the capital-labor ratio.
8. The calculation of the remaining variables is straightforward.

To match steady state moments, I run a minimization routine over the above steps, where the objective to be minimized is the Euclidean distance between model moments from their empirical targets.

Note that to allow for adjustment cost shocks I introduce a small alteration to the model in which steady adjustment are non-zero. In particular I define

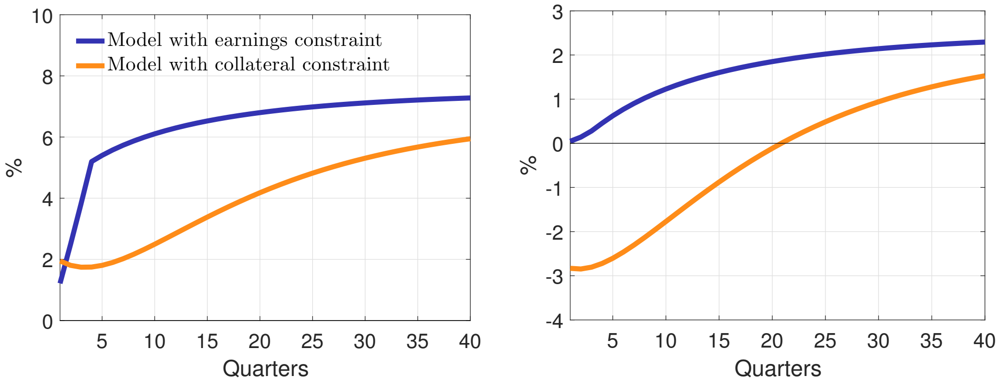
$$
\Phi_t\left(\frac{i_t}{i_{t-1}}\right) = \frac{\phi_t}{2}\left(\frac{i_t}{i_{t-1}} - \iota\right)^2,
$$

and set ι to 0.999.

IRF comparison with moving average earnings-based constraint

Figure 1.16: MODEL IRFS OF DEBT: MODIFIED EARNINGS-BASED CONSTRAINT

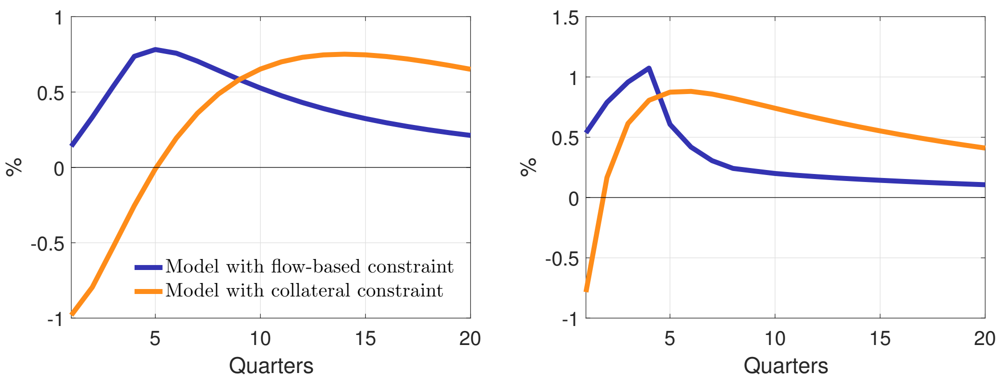
(a) Permanent TFP shock | (b) Permanent investment shock



Note: This figure repeats Figure 1.2 for a formulation of the earnings-based constraint in which current and three lags of earnings enter in equation (1.8). It displays the IRFs of firm debt to different shocks generated from the model, under the two alternative calibrations in which only the (in this case modified) earnings-based constraint (dark blue line) or only the collateral constraint (light orange line) is present. Panel (a) show the debt IRF to a positive TFP shock and Panel (b) to a positive investment shock. The structural parameters to generate these IRFs are shown in Table 1.2. I set $\rho_z = \rho_v = 1$, and $\sigma_z = \sigma_v = 0.05$.

Figure 1.17: MODEL IRFS TO INVESTMENT MARGIN SHOCKS: MODIFIED EARNINGS-BASED CONSTRAINT

(a) Persistent adjustment cost shock | (b) Persistent investment shock with $\overline{\phi} = 0$



Note: This figure repeats Figure 1.3 for a formulation of the earnings-based constraint in which current and three lags of earnings enter in equation (1.8). It displays IRFs of firm debt to different shocks generated from the model, under the two alternative calibrations in which only the (in this case modified) earnings-based constraint (dark blue line) or only the collateral constraint (light orange line) is present. Panel (a) plots the IRFs to an adjustment costs shock with $\rho_\phi = 0.5$ and $\sigma_\phi = 0.5$. Panel (b) repeats the investment shock IRFs from Figure 1.2 without the presence of investment adjustment costs ($\overline{\phi} = 0$).

Model IRFs of additional variables

Figure 1.18: IRFS TO PERMANENT TFP SHOCK

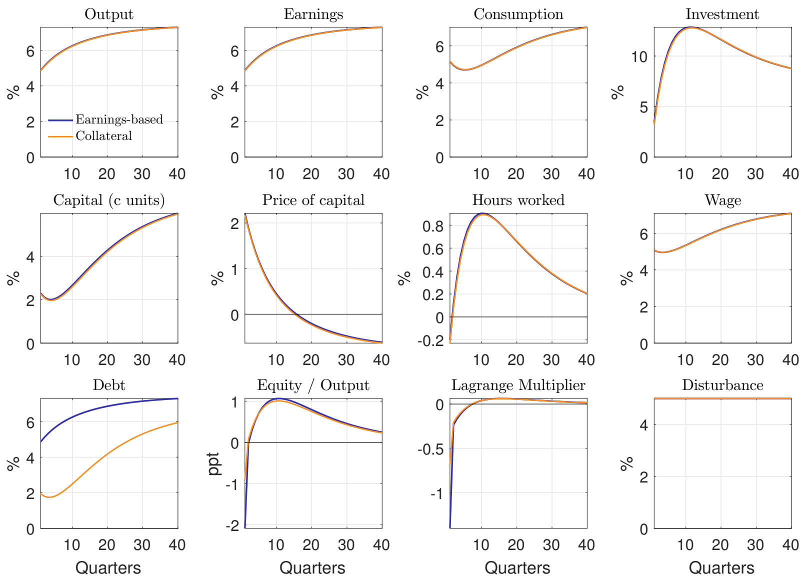
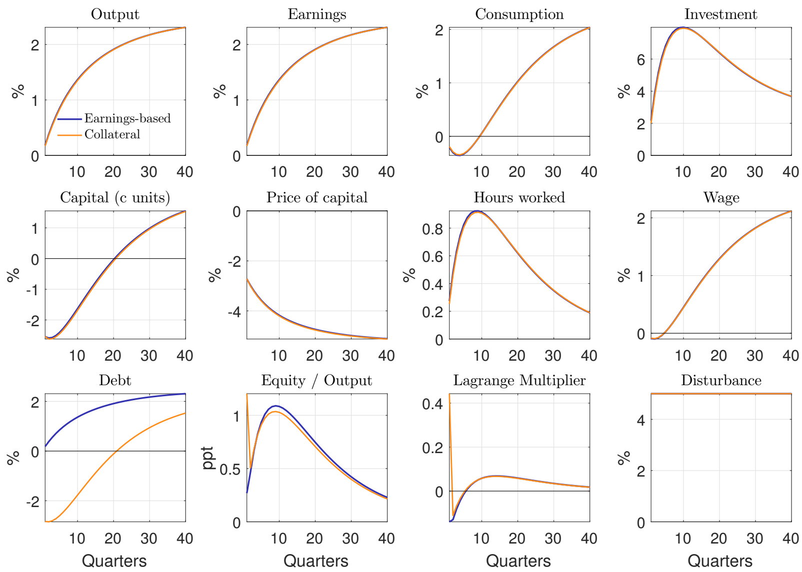


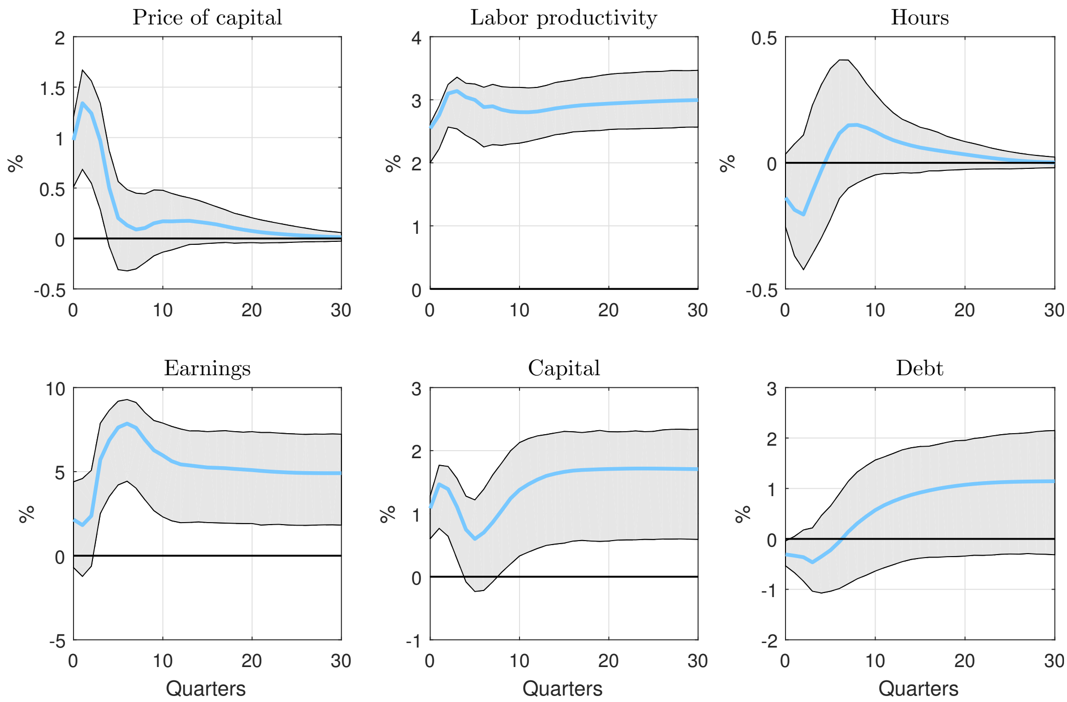
Figure 1.19: IRFS TO PERMANENT INVESTMENT SHOCK



1.7.5 Additional results for SVAR

IRFs to TFP shock

Figure 1.20: SVAR IRFS TO POSITIVE TFP SHOCK IDENTIFIED WITH LR RESTRICTIONS

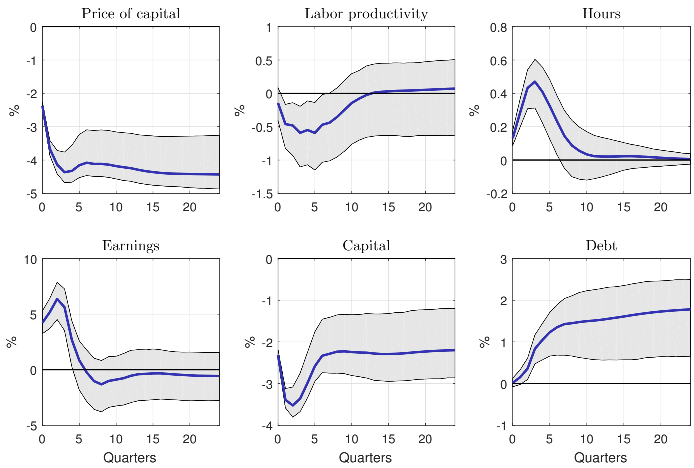


Note: The figure displays the IRFs to a TFP shock identified from an estimated SVAR model using US data. The identification scheme relies on long-run restrictions following Fisher (2006). The responses are shown for all six variables included in the system, in percent. The unit of the shock is one standard deviation. The sample period used for estimation is 1952:Q2 to 2016:Q4. 68% error bands are calculated using bootstrap techniques. This shock is identified using the same estimation procedure and identification scheme as the investment shock in the main text, but is not used to verify predictions from the theoretical macro model.

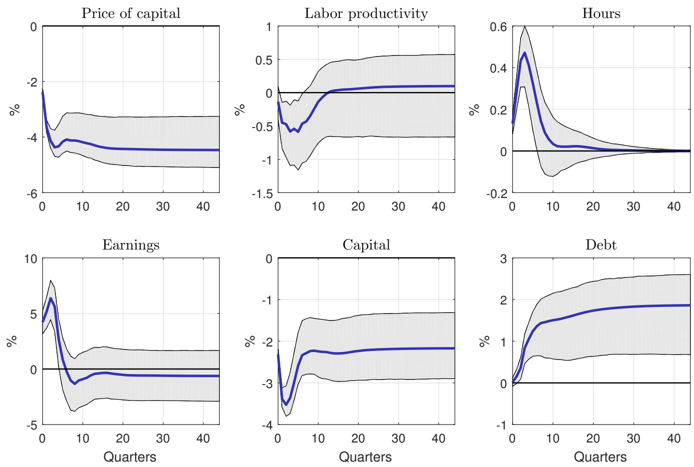
IRFs using medium-term restrictions

Figure 1.21: SVAR IRFS TO INVESTMENT SHOCK IDENTIFIED WITH MR RESTRICTIONS

(a) Identification based on 5-year horizon



(b) Identification based on 10-year horizon

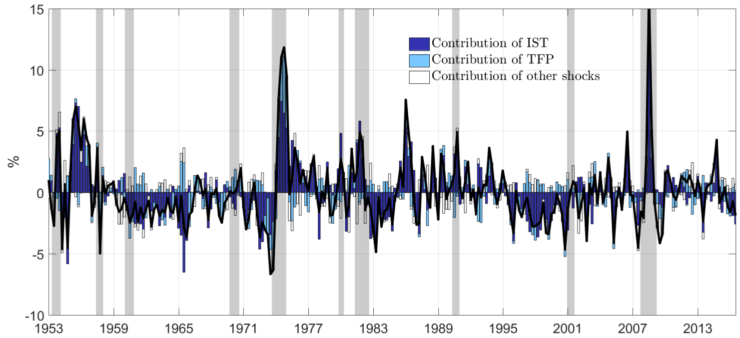


Note: The figure has the same scope as Figure 1.5 in the main text but uses a different identification scheme. This scheme is based on the method suggested by Barsky and Sims (2012). Panel (a) shows the results for a 5-year horizon ($h = 20$) and Panel (b) for a 10-year horizon ($h = 40$). In both cases, the responses are shown for all six variables included in the system, in percent. The unit of the shock is one standard deviation. The sample period used for estimation is 1952:Q2 to 2016:Q4. 68% error bands are calculated using bootstrap techniques. The figure shows a positive response of debt to an investment shock, which is in line with the predictions arising from a earnings-based borrowing constraint in the theoretical macro model.

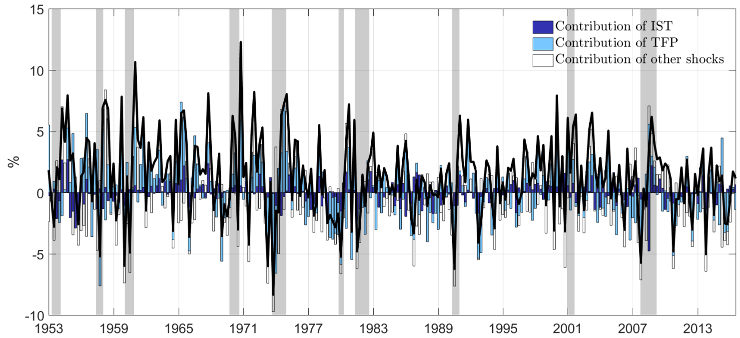
Historical decompositions for other variables

Figure 1.22: SVAR: HISTORICAL VARIANCE DECOMPOSITIONS

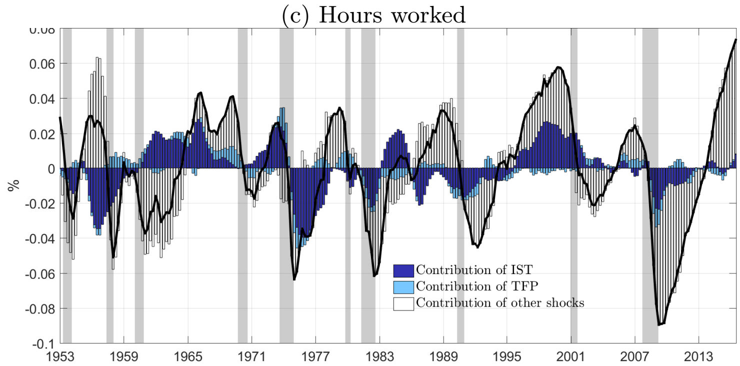
(a) Investment price



(b) Labor productivity



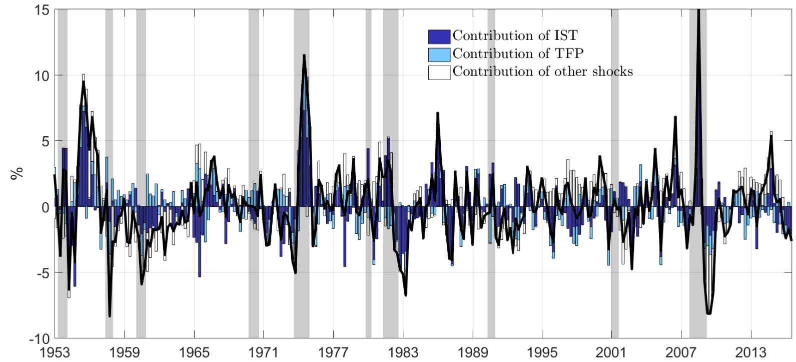
(c) Hours worked



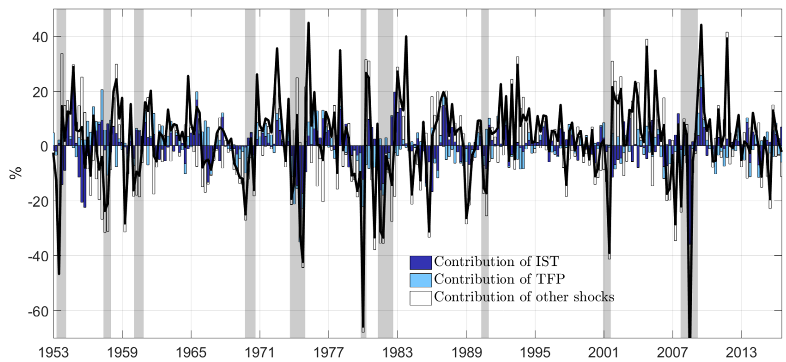
Note: Historical variance decomposition of variables as estimated by the SVAR model. The black line is the actual (detrended) data series. The bars indicate the contribution of different structural shocks to the variance of the respective observable as estimated by the SVAR model. The dark blue bars represent investment shocks, the light blue ones TFP shocks, and the contribution of shocks that remain unidentified are shown by the wight bars. Shaded areas indicate NBER recessions.

Figure 1.23: SVAR: HISTORICAL VARIANCE DECOMPOSITIONS

(a) Business earnings



(b) Capital stock



Note: Historical variance decomposition of variables as estimated by the SVAR model. The black line is the actual (detrended) data series. The bars indicate the contribution of different structural shocks to the variance of the respective observable as estimated by the SVAR model. The dark blue bars represent investment shocks, the light blue ones TFP shocks, and the contribution of shocks that remain unidentified are shown by the wight bars. Shaded areas indicate NBER recessions.

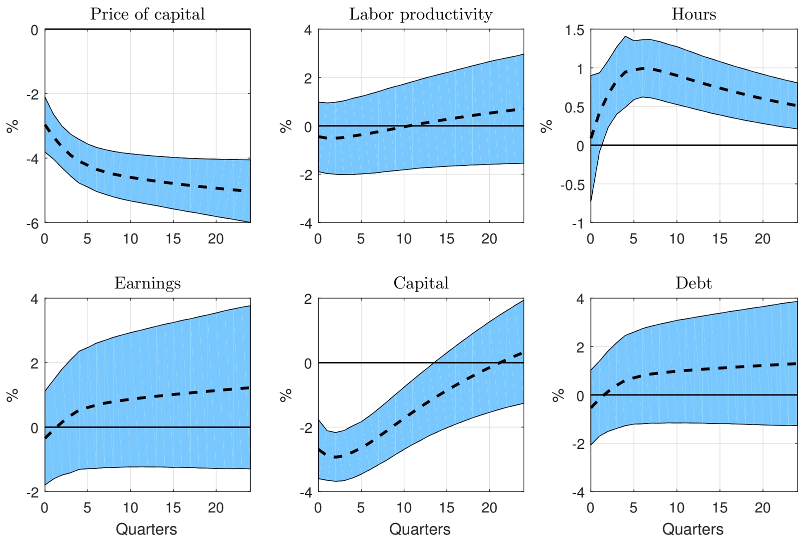
SVAR IRFs using simulated data

This appendix presents the results of a Monte Carlo exercise, which I set up as follows. I generate simulated data from the model in Section 1.3 and estimate the SVAR on this data. I repeatedly create two types of data samples, each generated from one of the two alternative borrowing constraint specifications (Panel (b) vs. Panel (c) in Table 1.2). I do so by randomly generating the time series in (1.15) from the model's solution. Specifically, I randomly draw permanent investment shocks, permanent TFP shocks, stationary government spending shocks (all with the same variance), and then plug them into the linearized policy rules of the model to generate observables. I then add iid measurement error to all series, calibrated to be 5% of the size of the structural shocks. For each sample type I generate 10,000 repetitions and run a SVAR identified with long-run restrictions on each of these samples. The identification procedure is carried out as described in the main text.

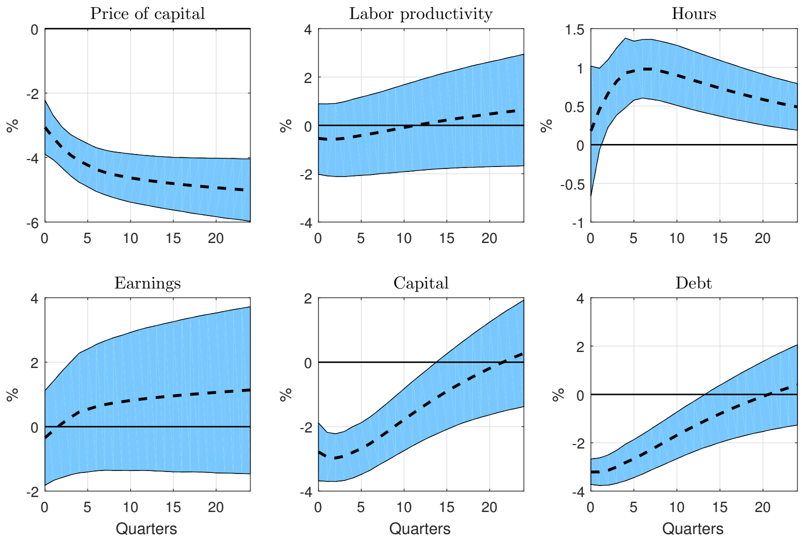
The results of this exercise are shown in Figure 1.24. Panel (a) plots the IRFs from estimations on samples generated with the earnings-based constraint, Panel (b) the equivalent with the collateral constraint. Each subpanel shows the mean (dashed line) and and 68% confidence sets (light blue area) across Monte Carlo repetitions. The figure shows that the direction of the debt IRF implied by the model is correctly picked up by the SVAR on average. Interestingly, while the negative debt response arising from the collateral constraint is estimated to be statistically significant, the positive one implied by the earnings constraint model is imprecisely estimated.

Figure 1.24: SVAR IRFS USING SIMULATED DATA

(a) SVAR IRFs to IST shock - Underlying data simulated with earnings-based constraint



(b) SVAR IRFs to IST shock - Underlying data simulated with collateral constraint



Note: The figure plots IRFs from an SVAR model estimated on data that is repeatedly simulated from the model in Section 1.3. Panel (a) uses the data generated with an earnings-based constraint, Panel (b) with a collateral constraint. In both cases, the data is generated from TFP shocks, investment shocks, an additional stationary demand shock. Normal iid measurement error is added to all series. 68% significance sets and means across 10,000 Monte Carlo repetitions are shown.

1.7.6 Additional results for firm-level projections

This appendix presents additional results on the estimation of equation (1.17) in Section 1.4.3 of the main text. Section 1.7.6 of the appendix reports the coefficient estimates of the difference between earnings and collateral borrowers' debt IRFs, which serves as a formal test of the difference between the IRFs shown in Figure 1.8. Section 1.7.6 shows the results of Figure 1.8 for an alternative specification in which I estimate the IRF to a fall in the relative price of equipment investment, instrumented by the investment shock (rather than the debt response to the investment shock directly). Section 1.7.6 contains the results for a firm fixed effects regression specification. In Section 1.7.6, the main results displayed in Figure 1.8 are shown also for the two additional groups, which are firms subject to both covenants and collateral, as well as firms that are subject to neither.

Significance of the difference between heterogeneous IRFs

Table 1.12: ESTIMATES OF THE DIFFERENCE BETWEEN IRF COEFFICIENTS

	Classification based on specific assets	**Classification based on secured revolvers**
$\beta_0^{earn} - \beta_0^{coll}$	0.0328	-0.0029
	(0.0213)	(0.0248)
$\beta_1^{earn} - \beta_1^{coll}$	0.0308	0.0004
	(0.0318)	(0.0285)
$\beta_2^{earn} - \beta_2^{coll}$	0.0340	0.0162
	(0.0282)	(0.0307)
$\beta_3^{earn} - \beta_3^{coll}$	0.0511	0.0511
	(0.0334)	(0.0365)
$\beta_4^{earn} - \beta_4^{coll}$	0.0600*	0.0464
	(0.0345)	(0.0404)
$\beta_5^{earn} - \beta_5^{coll}$	0.0491	0.0384
	(0.0331)	(0.0370)
$\beta_6^{earn} - \beta_6^{coll}$	0.0581*	0.0400
	(0.0351)	(0.0395)
$\beta_7^{earn} - \beta_7^{coll}$	0.0688*	0.0642*
	(0.0353)	(0.0356)
$\beta_8^{earn} - \beta_8^{coll}$	0.0865**	0.0813**
	(0.0355)	(0.0358)
$\beta_9^{earn} - \beta_9^{coll}$	0.0810**	0.0725*
	(0.0389)	(0.0386)
$\beta_{10}^{earn} - \beta_{10}^{coll}$	0.0773*	0.0624
	(0.0406)	(0.0403)
$\beta_{11}^{earn} - \beta_{11}^{coll}$	0.0927**	0.0893**
	(0.0420)	(0.0432)
$\beta_{12}^{earn} - \beta_{12}^{coll}$	0.0690	0.0658
	(0.0433)	(0.0442)

Note: The table shows estimates of the difference between the IRF of earnings borrowers and collateral borrowers as estimated by equation (1.17) in the main text. The left column shows these estimates for the specification corresponding to Panel (a) of Figure 1.8 and the right column for Panel (b). Standard errors are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

IV strategy

The results presented here study the responses of firm debt to a fall in the relative price of investment goods, instrumented by the exogenous investment shock, rather than considering the direct responses to the shock itself, as formulated by equation (1.17) and presented in the main text. To this end, equation (1.17) from the main text is modified to

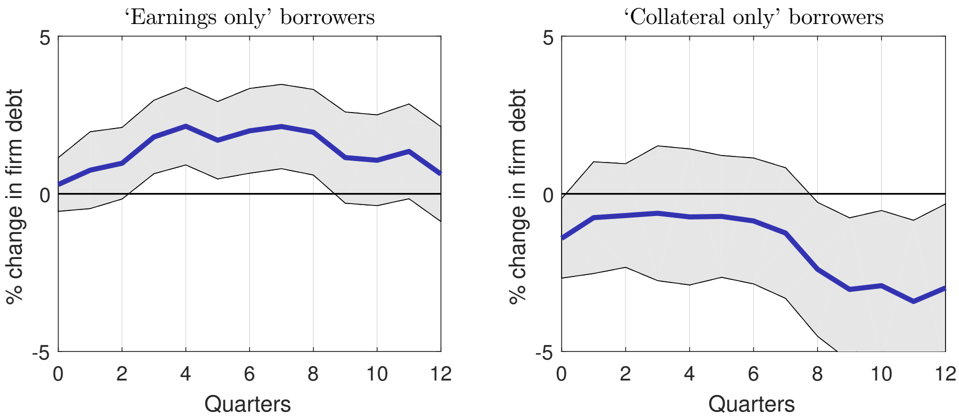
$$\begin{aligned} log(b_{i,t+h}) = \alpha_h + \beta_h p_{k,t} + \boldsymbol{\gamma X_{i,t}} \\ + \beta_h^{earn} \mathbb{1}_{i,t,earn} \times p_{k,t} + \alpha_h^{earn} \mathbb{1}_{i,t,earn} \\ + \beta_h^{coll} \mathbb{1}_{i,t,coll} \times p_{k,t} + \alpha_h^{coll} \mathbb{1}_{i,t,coll} + \gamma t + \eta_{i,t+h}, \end{aligned} \tag{1.41}$$

where $p_{k,t}$ is defined as in Section 1.4.1. Equation (1.41) is then estimated by using $\hat{u}_{IST,t}$ as an IV for $p_{k,t}$.[77] The results for this specification, presented analogous to Figure 1.8, are shown in Figure 1.25 below. They paint a very similar picture to the results in the main text. The responses are smaller in magnitude, and standard errors are lower relative to when the shock is used as a regressor directly.

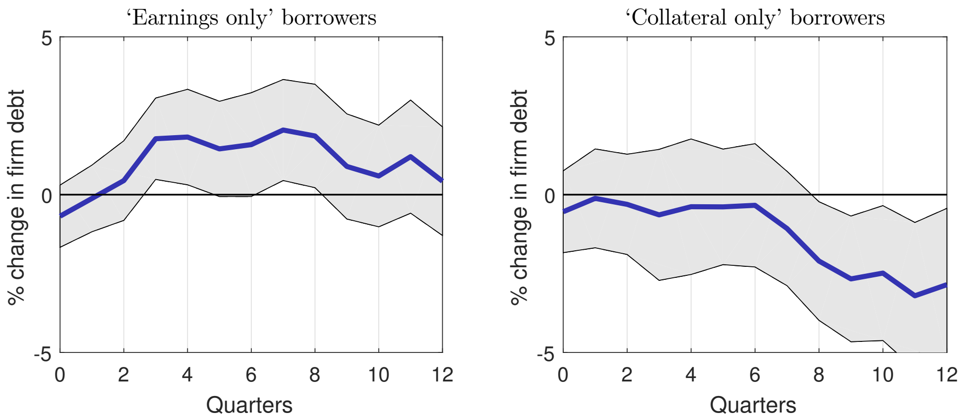
[77] See Jordà, Schularick, and Taylor (2017) for a similar approach in different context.

Figure 1.25: FIRM-LEVEL IRFS TO FALL IN INVESTMENT PRICE, IST SHOCK AS IV

(a) Using collateral classification based on specific assets



(b) Using collateral classification based on secured revolvers

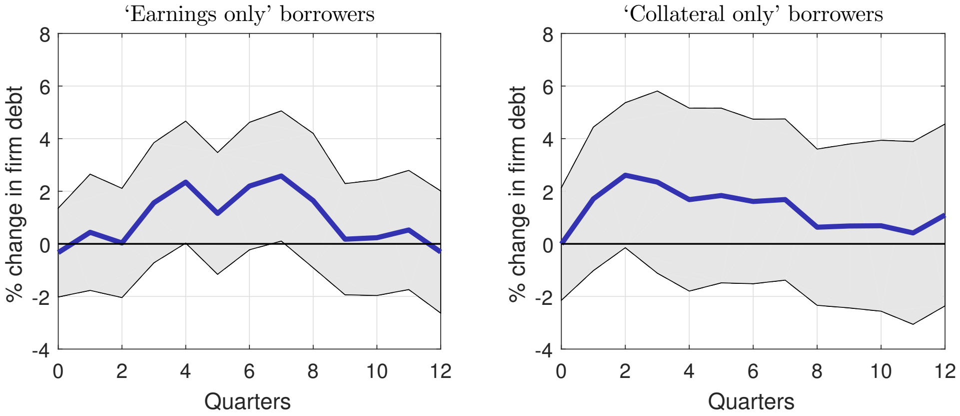


Note: This figure repeats Figure 1.8 from the text but instead plots the IRFs to a fall in the relative price of investment, instrumented with the investment shock, see equation (1.25) above. In both panels of the figure, the debt IRF for borrowers with earnings covenants and no collateral (left) and borrowers without earnings covenants but with collateral (right) are plotted. The results are based on a specification with detailed firm-level controls (3-digit industry fixed effects, size as measured by number of employees, growth of real sales and other macroeconomic shocks). Panel (a) uses the collateral classification based on whether a loan is backed by specific assets or not (see details in Section 1.2). Panel (b) uses an alternative grouping where secured revolvers are categorized as collateralized debt (see Lian and Ma, 2018). The investment shock is identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information, with balance sheet variables from the Compustat quarterly data base. 90% bands are calculated using standard errors clustered at the 3-digit industry level. The IRFs shown in the figure are consistent with the model's prediction of a positive debt response under an earnings-based constraint and a negative one under a collateral constraint.

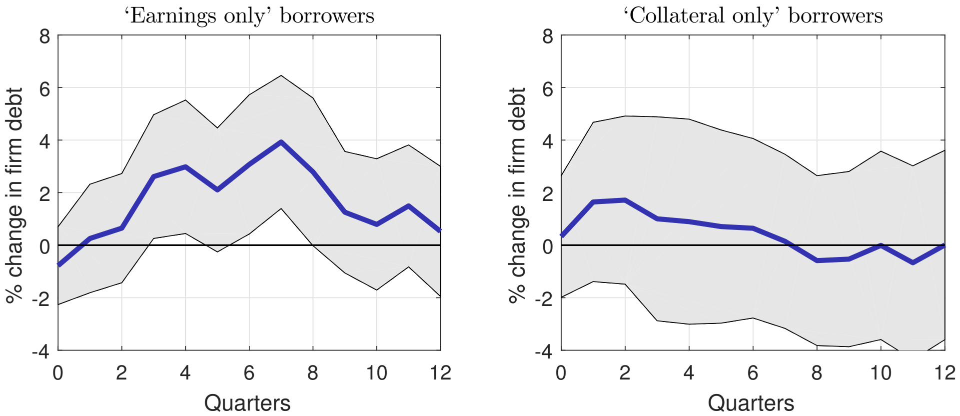
Results for specification with firm fixed effects

Figure 1.26: FIRM-LEVEL IRFS INVESTMENT SHOCK: FIRM FIXED EFFECTS SPECIFICATION

(a) Using collateral classification based on specific assets



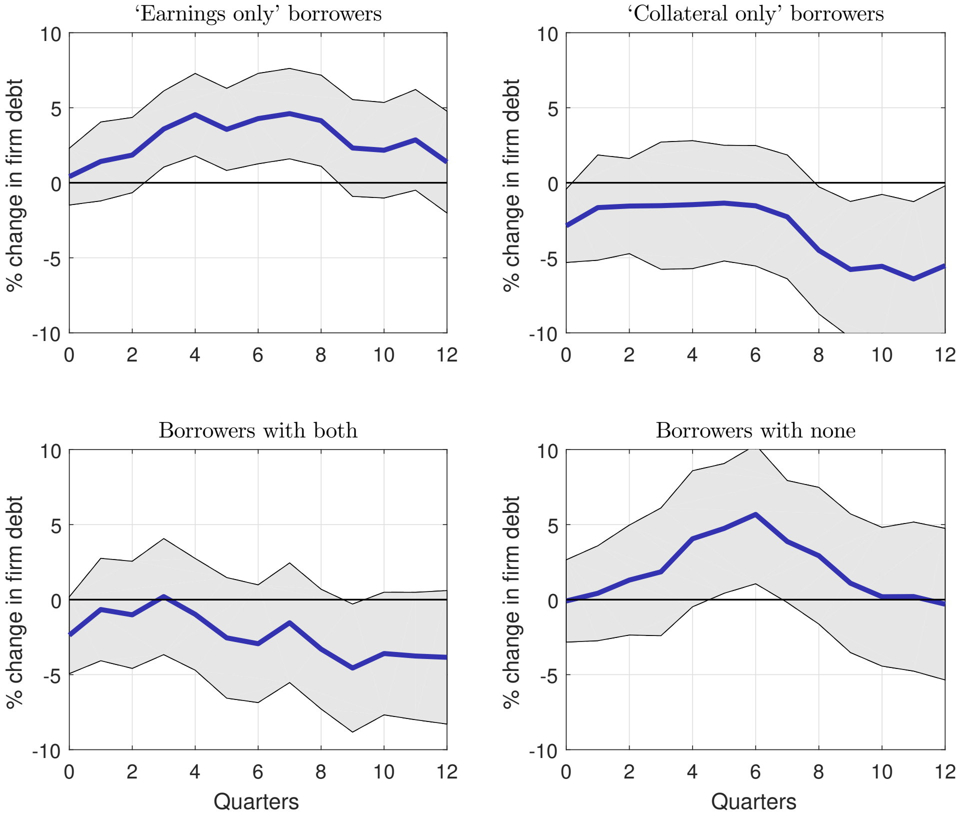
(b) Using collateral classification based on secured revolvers



Note: This figure repeats Figure 1.8 from the text for a regression specification with firm-fixed effects. The figure displays average IRFs of firm borrowing for different firm groups, estimated using the method of Jordà (2005) in a panel data context, see equation (1.17). In both panels of the figure, the debt IRF for borrowers with earnings covenants and no collateral (left) and borrowers without earnings covenants but with collateral (right) are plotted. Panel (a) uses the collateral classification based on whether a loan is backed by specific assets or not (see details in Section 1.2). Panel (b) uses an alternative grouping where secured revolvers are categorized as collateralized debt (see Lian and Ma, 2018). The investment shock is identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information, with balance sheet variables from the Compustat quarterly data base. 90% bands are calculated using standard errors clustered at the 3-digit industry level. The IRFs shown in the figure are consistent with the model's prediction of a positive debt response under an earnings-based constraint and a negative one under a collateral constraint.

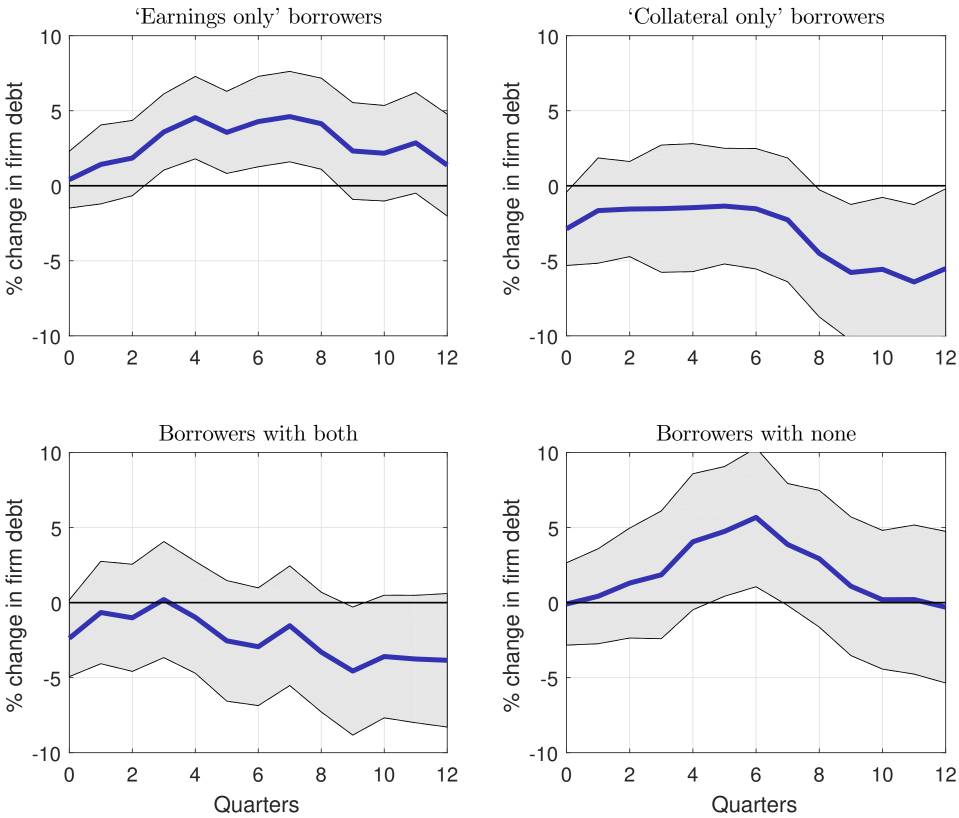
Results for all four firm groups

Figure 1.27: IRFS FOR ALL FOUR CATEGORIES: COLLATERAL CLASSIFICATION BASED ON SPECIFIC ASSETS



Note: This figure repeats Panel (a) of Figure 1.8 in the main text, and additionally plots the IRFs of the remaining two firm groups: borrowers with both earnings covenants and collateral, and borrowers with neither. The results are based on a specification with detailed firm-level controls (3-digit industry fixed effects, size as measured by number of employees, growth of real sales and other macroeconomic shocks). The investment shock is identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information, with balance sheet variables from the Compustat quarterly data base. 90% bands are calculated using standard errors clustered at the 3-digit industry level.

Figure 1.28: IRFS FOR ALL FOUR CATEGORIES: COLLATERAL CLASSIFICATION BASED ON SECURED REVOLVERS



Note: This figure repeats Panel (b) of Figure 1.8 in the main text, and additionally plots the IRFs of the remaining two firm groups: borrowers with both earnings covenants and collateral, and borrowers with neither. The results are based on a specification with detailed firm-level controls (3-digit industry fixed effects, size as measured by number of employees, growth of real sales and other macroeconomic shocks). The investment shock is identified using the SVAR model in the previous section, based on long-run restrictions following Fisher (2006). The data set used is a merge of Dealscan loan-level information, with balance sheet variables from the Compustat quarterly data base. 90% bands are calculated using standard errors clustered at the 3-digit industry level.

1.7.7 Details on the quantitative model of Section 1.5

Model setup

The model is a variant of the medium scale New Keynesian model introduced by Smets and Wouters (2007), similar to Jermann and Quadrini (2012). The core of the model is that of Section 1.3 but a variety of additional frictions are added.

Final good firm The final good firm produces a consumption good Y_t using inputs $y_{i,t}$ that are provided by intermediate producers. The production function is

$$Y_t = \left(\int_0^1 y_{i,t}^{\frac{1}{\eta_t}} \, di \right)^{\eta_t}. \tag{1.42}$$

η_t is a stochastic price markup disturbance. The final good is sold to households at price P_t and intermediate inputs are purchased at price $p_{i,t}$. The optimality conditions of the final good firm can be written as

$$p_{i,t} = P_t Y_t^{\frac{\eta_t - 1}{\eta_t}} y_{i,t}^{\frac{1-\eta_t}{\eta_t}} \tag{1.43}$$

which is the demand function that intermediate producers take as given, and intermediate prices aggregate to the economy's price level as $P_t = \left(\int_0^1 p_{i,t}^{\frac{1}{1-\eta_t}} \, di \right)^{1-\eta_t}$.

Intermediate goods firms There is a continuum of size 1 of firms, which produce an intermediate good $y_{i,t}$ that is sold at price $p_{i,t}$ to a final good producer. The production of intermediate goods is based on a Cobb-Douglas production function

$$y_{i,t} = z_t (u_{i,t} k_{i,t-1})^{\alpha} n_{i,t}^{1-\alpha}, \tag{1.44}$$

where TFP, z_t, is common across firms and will be subject to stochastic shocks. $k_{i,t-1}$ is capital, which is owned and accumulated by firms and predetermined at the beginning of the period. $u_{i,t}$ is the utilization rate of capital, which is an endogenous choice taken subject to a cost to be specified further below. $\alpha \in (0, 1)$ is the capital share in production. $n_{i,t}$ denotes labor used by firm i at the wage rate $w_{i,t}$, which is a composite of different labor types j that will be supplied by households:

$$n_{i,t} = \left(\int_0^1 n_{j,i,t}^{\frac{1}{\vartheta_t}} \, dj \right)^{\vartheta_t}, \tag{1.45}$$

where v_t is stochastic shock that affects demand for labor. A firm's period earnings flow, or operational profits, is denoted as $\pi_{i,t}$ and defined as

$$\pi_{i,t} \equiv y_{i,t} - w_{i,t} n_{i,t}. \tag{1.46}$$

As in the model in Section 1.3 the law of motion of capital is

$$k_{i,t} = (1-\delta)k_{i,t-1} + v_t \left[1 - \frac{\phi}{2}\left(\frac{i_{i,t}}{i_{i,t-1}}\right)^2\right] i_{i,t}. \tag{1.47}$$

MEI shocks enter via the disturbance v_t. Note that in the quantitative application I do not allow for shocks to ϕ_t for comparability with previous studies.

Firms take (1.43) as given given when setting their price. Combining this equation with the production function, the price can be written as a function of aggregate variables and individual inputs, so that

$$p_{i,t} = P_t Y_t^{\frac{\eta_t - 1}{\eta_t}} \left(z_t (u_{i,t} k_{i,t-1})^{\alpha} n_{i,t}^{1-\alpha}\right)^{\frac{1-\eta_t}{\eta_t}}. \tag{1.48}$$

The capital utilization cost is specified as

$$\Xi(u_t) = \xi_1 (u_t^{1+\xi_2} - 1)/(1-\xi_2) \tag{1.49}$$

The parameter ξ_1 is calibrated to generate steady state utilization of 1.

The firm sets prices subject to a Rotemberg adjustment cost. As discussed in detail by Jermann and Quadrini (2012), this approach to generating price rigidities – as opposed to, say Calvo pricing – substantially facilitates the aggregation of the decision of individual firms when financial frictions are introduced.

Specifically, a firm that has previously set price $p_{i,t-1}$ faces adjustment costs

$$\tilde{\Phi}(p_{i,t-1}, p_{i,t}, Y_t) = \frac{\tilde{\phi}}{2}\left(\frac{p_{i,t}}{p_{i,t-1}} - 1\right)^2 Y_t. \tag{1.50}$$

The firm has access to debt, which is limited by weighting between an earnings-based and a collateral component. The details of this constraints are given in the main text, see the description of equation (1.18).

Firm maximization problem. The objective of firms is similar to what is descried in equation (1.11) in the more stylized model of Section 1.3. In the New Keynesian setting, firms maximize the flow of (nominal) dividends, discounted with the household's stochastic discount factor, subject the flow of dividends equation (which now contains also price adjustment and utilization costs), the borrowing constraint (1.18), the law of motion of capital (1.47) and the demand function given by (1.43). They now also choose their price $p_{i,t}$ and utilization rate $u_{i,t}$, in addition to $d_{i,t}$, $n_{i,t}$, $i_{i,t}$, $k_{i,t}$, and $b_{i,t}$.

Households There is a continuum of size 1 of households. Household $j's$ expected lifetime utility is given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \gamma_t \beta^t \left(\frac{(c_{j,t} - hc_{j,t-1})^{1-\sigma}}{1-\sigma} - \chi \frac{n_{j,t}^{1+\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}}\right) \tag{1.51}$$

where γ_t is a preference disturbance and h captures external consumption habits. The parameter ϵ denotes the elasticity of labor supply. Households supply individual labor types $n_{j,t}$ and charge wage rate $w_{j,t}$. The budget constraint is

$$c_{j,t} + \frac{b_{j,t}}{1+r_t} + p_t^f s_{j,t} + T_{j,t} + \int q_{j,t+1}^{\bar{\omega}} a_{j,t+1} dw_{j,t} = w_{j,t} n_{j,t} + b_{j,t-1} + P_t d_{j,t} + p_t^f s_{j,t-1}. \tag{1.52}$$

$a_{j,t+1}$ are holdings of state-contingent claims with which households can insure against wage shocks. They are traded at price $q_{j,t+1}^w$. The notation in (1.52) is otherwise similar as in the stylized mode of Section 1.3.

The demand for labor coming from the intermediate goods firms is given by

$$n_{j,t} = \left(\frac{w_{j,t}}{W_t}\right)^{-\frac{\vartheta_t}{\vartheta_t - 1}} n_t, \tag{1.53}$$

where W_t and n_t are the aggregate wage and employment level, respectively. (1.53) is taken as given by the household when choosing $n_{j,t}$ and $w_{j,t}$.

Households face wage rigidities, which arise, in the spirit of Calvo, from the fact that a given firm can only change their wage with probability $(1 - \bar{\omega})$. From the optimization problem I derive a log-linear optimal wage equation. Given that all households make the same choices, this implies a sluggish low of motion for the aggregate wage rate W_t (for details, see Jermann and Quadrini, 2012).

Household's optimality condition for bonds implies an Euler euqation in which the real return $(1+r_t)\left(\frac{P_t}{P_{t+1}}\right)$ is priced with the stochastic discount factor $SDF_{t,t+1} \equiv \frac{\Lambda_{t+1}}{\Lambda_t} = \frac{\beta\gamma_{t+1} u_{c_{t+1}}}{\gamma_t u_{c_t}}$, where $u(\cdot)$ denotes the period utility function in (1.51).

Government The government's budget constraint, in nominal terms, reads

$$T_t = \frac{b_t}{R_t} - \frac{b_t}{(1+r_t)} + P_t G_t, \tag{1.54}$$

where T_t are nominal lump sum taxes levied on households, the term $\frac{b_t}{R_t} - \frac{b_{k,t}}{(1+r_{k,t})}$ is the tax subsidy given to firms, and G_t is a real spending shock that follows an exogenous stochastic process.

Monetary policy There is a Taylor rule specified as

$$\frac{1+r_t}{1+\bar{r}} = \left[\frac{1+r_{t-1}}{1+\bar{r}}\right]^{\rho_R} \left[\left(\frac{\pi_t^p}{\bar{\pi}^p}\right)^{\nu_1} \left(\frac{Y_t}{Y_{t-1}}\right)^{\nu_2}\right]^{1-\rho_R} \left[\frac{Y_t/Y_t^*}{Y_{t-1}/Y_{t-1}^*}\right]^{\nu_3} \varsigma_t, \tag{1.55}$$

such that interest rates react to deviations of inflation from steady state, output growth, and output growth in deviations from it steady state.[78] Note that I denote inflation by π_t^p, not to be confused with firm profits $\pi_{i,t}$. $\rho_R > 0$ captures interest

[78] See Jermann and Quadrini (2012) for more details.

rate smoothing. ς_t is a stochastic disturbance that captures monetary policy shocks.

Stochastic processes The model features eight structural disturbances, capturing shocks to TFP, investment, preferences, price markups, wage markups, fiscal policy, monetary policy and financial conditions. The follow autoregressive processes of order one:

$$\begin{aligned}
log(z_t) &= (1-\rho_z)log(\overline{z}) + \rho_z log(z_{t-1}) + u_{z,t} && (1.56)\\
log(v_t) &= (1-\rho_v)log(\overline{v}) + \rho_v log(v_{t-1}) + u_{v,t} && (1.57)\\
log(\gamma_t) &= (1-\rho_\gamma)log(\overline{\gamma}) + \rho_\gamma log(\gamma_{t-1}) + u_{\gamma,t} && (1.58)\\
log(\eta_t) &= (1-\rho_\eta)log(\overline{\eta}) + \rho_\eta log(\eta_{t-1}) + u_{\eta,t} && (1.59)\\
log(\vartheta_t) &= (1-\rho_\vartheta)log(\overline{\vartheta}) + \rho_\vartheta log(\vartheta_{t-1}) + u_{\vartheta,t} && (1.60)\\
log(g_t) &= (1-\rho_g)log(\overline{g}) + \rho_g log(g_{t-1}) + u_{g,t} && (1.61)\\
log(\varsigma_t) &= (1-\rho_\varsigma)log(\overline{\varsigma}) + \rho_\varsigma log(\varsigma_{t-1}) + u_{\varsigma,t} && (1.62)\\
log(\xi_t) &= (1-\rho_\xi)log(\overline{\xi}) + \rho_\xi log(\xi_{t-1}) + u_{\xi,t} && (1.63)
\end{aligned}$$

The error terms follow standard deviations $\{\sigma_z, \sigma_v, \sigma_\gamma, \sigma_\eta.\sigma_\vartheta, \sigma_G, \sigma_\varsigma, \sigma_\xi\}$ I normalize $\bar{z} = \bar{v} = \bar{\gamma} = \bar{\varsigma} = \bar{\xi} = 1$, calibrate $\bar{g}$ to match the US purchases-to-output ratio, and estimate $\bar{\eta}$ and $\bar{\vartheta}$.

Additional estimation results

Table 1.13: PRIORS AND POSTERIORS FROM BAYESIAN ESTIMATION

	Prior shape	Prior Mean	Prior Std	Post. mean	90% HPD interval	
ω	Uniform	0.5	$\sqrt{12}$	0.8992	0.7968	1
$\tilde{\phi}$	Inv-Gamma	0.1	0.3	23.433	19.7002	26.9731
σ	Normal	1.5	0.37	1.6721	1.1921	2.1902
ϵ	Normal	2	0.75	1.6229	0.6464	2.5399
h	Beta	0.5	0.15	0.9709	0.9615	0.981
$\bar{\omega}$	Beta	0.5	0.15	0.843	0.7718	0.9049
ϕ	Inv-Gamma	0.1	0.3	5.7251	4.2177	7.4163
ψ	Beta	0.5	0.15	0.4556	0.2548	0.658
κ	Inv-Gamma	0.2	0.1	0.3395	0.2072	0.4685
ρ_R	Beta	0.75	0.1	0.6434	0.5823	0.703
ν_1	Normal	1.5	0.25	2.3287	1.9708	2.6836
ν_2	Normal	0.12	0.05	-0.0487	-0.0862	-0.0117
ν_3	Normal	0.12	0.05	0.2181	0.1536	0.2865
$\bar{\eta}$	Beta	1.2	0.1	1.459	1.3817	1.5376
$\bar{\vartheta}$	Beta	1.2	0.1	1.1519	1.0229	1.2823
ρ_z	Beta	0.5	0.2	0.9903	0.9844	0.9967
ρ_{gz}	Beta	0.5	0.2	0.9328	0.8756	0.9908
ρ_v	Beta	0.5	0.2	0.754	0.6814	0.8303
ρ_γ	Beta	0.5	0.2	0.3669	0.2594	0.4754
ρ_η	Beta	0.5	0.2	0.8893	0.8407	0.9364
ρ_ϑ	Beta	0.5	0.2	0.3034	0.2215	0.3831
ρ_G	Beta	0.5	0.2	0.9442	0.9084	0.9813
ρ_ς	Beta	0.5	0.2	0.4173	0.3167	0.5171
ρ_ξ	Beta	0.5	0.2	0.9893	0.9821	0.9967
σ_z	Inv-Gamma	0.001	0.05	0.0073	0.0067	0.0079
σ_v	Inv-Gamma	0.001	0.05	0.1343	0.0933	0.1754
σ_γ	Inv-Gamma	0.001	0.05	0.247	0.1786	0.3092
σ_η	Inv-Gamma	0.001	0.05	0.0162	0.0133	0.0191
σ_ϑ	Inv-Gamma	0.001	0.05	1.3028	0.9642	1.5692
σ_G	Inv-Gamma	0.001	0.05	0.0176	0.0161	0.019
σ_ς	Inv-Gamma	0.001	0.05	0.0094	0.0086	0.0102
σ_ξ	Inv-Gamma	0.001	0.05	0.0395	0.0356	0.0434

Table 1.14: PRIORS AND POSTERIORS FOR MODEL WITHOUT BORROWING CONSTRAINT

	Prior shape	**Prior Mean**	**Prior Std**	**Post. mean**	**90% HPD interval**	
$\tilde{\phi}$	Inv-Gamma	0.1	0.3	7.083	6.26	7.9168
σ	Normal	1.5	0.37	2.0603	1.7621	2.3992
ϵ	Normal	2	0.75	0.7779	0.6112	0.9231
h	Beta	0.5	0.15	0.8091	0.7676	0.8505
$\bar{\omega}$	Beta	0.5	0.15	0.2911	0.199	0.424
ϕ	Inv-Gamma	0.1	0.3	5.8699	5.0646	6.4816
ψ	Beta	0.5	0.15	0.7283	0.625	0.8244
κ	Inv-Gamma	0.2	0.1	0.8349	0.6237	1.0961
ρ_R	Beta	0.75	0.1	0.654	0.6256	0.6834
ν_1	Normal	1.5	0.25	3.0458	2.9881	3.0903
ν_2	Normal	0.12	0.05	-0.0166	-0.0539	0.0156
ν_3	Normal	0.12	0.05	0.1713	0.1063	0.2213
$\bar{\eta}$	Beta	1.2	0.1	1.4217	1.3582	1.4741
$\bar{\vartheta}$	Beta	1.2	0.1	1.0774	1.0462	1.1093
ρ_z	Beta	0.5	0.2	0.983	0.9749	0.9911
ρ_{gz}	Beta	0.5	0.2	0.9509	0.9113	0.9928
ρ_v	Beta	0.5	0.2	0.7128	0.6586	0.7668
ρ_γ	Beta	0.5	0.2	0.8122	0.7699	0.8559
ρ_η	Beta	0.5	0.2	0.9775	0.9606	0.9957
ρ_ϑ	Beta	0.5	0.2	0.9848	0.9734	0.9965
ρ_G	Beta	0.5	0.2	0.9781	0.9612	0.996
ρ_ς	Beta	0.5	0.2	0.1896	0.1158	0.2653
σ_z	Inv-Gamma	0.001	0.05	0.0072	0.0067	0.0077
σ_v	Inv-Gamma	0.001	0.05	0.1865	0.1709	0.2027
σ_γ	Inv-Gamma	0.001	0.05	0.0585	0.0505	0.0663
σ_η	Inv-Gamma	0.001	0.05	0.0113	0.01	0.0124
σ_ϑ	Inv-Gamma	0.001	0.05	0.0628	0.0521	0.0727
σ_G	Inv-Gamma	0.001	0.05	0.0177	0.0162	0.0191
σ_ς	Inv-Gamma	0.001	0.05	0.0097	0.0088	0.0106

Table 1.15: VARIANCE DECOMPOSITION WITHOUT BORROWING CONSTRAINTS (%)

	TFP	**Inv**	**Pref**	**Price**	**Wage**	**Gov**	**Mon**	**Fin**
Output growth	11.9	27.4	1.5	11.7	33.1	10.2	4.2	-
Consumption growth	11.1	6.6	14.7	7.6	52.7	0.9	6.3	-
Investment growth	1.7	82.1	7.7	1.5	6.8	0.0	0.3	-
Inflation	12.4	12.1	34.2	17.5	9.8	0.3	13.7	-
Interest rate	6.0	18.2	38.9	8.2	6.9	0.5	21.3	-
Employment growth	25.7	25.3	0.6	3.4	33.0	9.3	2.8	-
Wage growth	27.4	3.2	5.4	55.3	6.2	0.1	2.4	-
Debt issuance	0.37	0.29	28.66	4.44	65.26	0.52	0.46	-

Note: Repeats Table 1.3 from the text, but for a version of the model that is re-estimated without any borrowing constraints. In this model, debt is in zero net supply so debt issuance cannot be used as an observable, so the financial shock is dropped. The table shows the infinite horizon forecast error variance decomposition of the observables used for this model version. The decompositions are calculated at the estimated posterior means. Each row presents the decomposition for a given observable, columns correspond to different structural shocks that feature in the model: TFP-Total productivity shock; Inv-Investment shock; Pref-Preference shock; Price-Price markup shock; Wage-Wage markup shock; Gov-Government spending shock; Mon-Monetary policy shock; Fin-Financial shock.

1.8 Bibliography

Acharya, V. V., H. Almeida, and M. Campello (2013): "Aggregate Risk and the Choice between Cash and Lines of Credit," *The Journal of Finance*, 68, 2059–2116.

Adler, K. (2018): "Financial Covenants, Firm Financing and Investment," *Unpublished Manuscript.*

Aghion, P. and P. Bolton (1992): "An incomplete contracts approach to financial contracting," *The review of economic Studies*, 59, 473–494.

Altinkilic, O. and R. Hansen (2000): "Are There Economies of Scale in Underwriting Fees? Evidence of Rising External Financing Costs," *Review of Financial Studies*, 13, 191–218.

An, S. and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26, 113–172.

Angeletos, G.-M., F. Collard, and H. Dellas (2018): "Business Cycle Anatomy," Working Paper 24875, National Bureau of Economic Research.

Arellano, C. and E. G. Mendoza (2002): "Credit Frictions and 'Sudden Stops' in Small Open Economies: An Equilibrium Business Cycle Framework for Emerging Markets Crises," NBER Working Papers 8880, National Bureau of Economic Research, Inc.

Azariadis, C., L. Kaas, and Y. Wen (2016): "Self-Fulfilling Credit Cycles," *The Review of Economic Studies*, 83, 1364–1405.

Bahaj, S., A. Foulis, G. Pinter, and P. Surico (2018): "The Employment Effects of Monetary Policy: Evidence from Firm-level data," Working paper.

Barsky, R. B. and E. R. Sims (2012): "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, 102, 1343–77.

Bates, T. W., K. M. Kahle, and R. M. Stulz (2009): "Why Do U.S. Firms Hold So Much More Cash than They Used To?" *The Journal of Finance*, 64, 1985–2021.

Ben Zeev, N. and H. Khan (2015): "Investment-Specific News Shocks and US Business Cycles," *Journal of Money, Credit and Banking*, 47, 1443–1464.

Bernanke, B. and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review*, 79, 14–31.

BERNANKE, B. S. AND M. GERTLER (1995): "Inside the Black Box: The Credit Channel of Monetary Policy Transmission," *The Journal of Economic Perspectives*, 9, 27–48.

BERNANKE, B. S., M. GERTLER, AND S. GILCHRIST (1999): "The financial accelerator in a quantitative business cycle framework," *Handbook of Macroeconomics*, 1, 1341 – 1393.

BIANCHI, J. (2011): "Overborrowing and Systemic Externalities in the Business Cycle," *American Economic Review*, 101, 3400–3426.

BLANCHARD, O. J. AND D. QUAH (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

BRADLEY, M. AND M. R. ROBERTS (2015): "The Structure and Pricing of Corporate Debt Covenants," *Quarterly Journal of Finance*, 05, 1550001.

BROOKS, W. AND A. DOVIS (2018): "Credit market frictions and trade liberalization," *Working paper*.

BUERA, F. J. AND B. MOLL (2015): "Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity," *American Economic Journal: Macroeconomics*, 7, 1–42.

CAO, D. AND G. NIE (2017): "Amplification and Asymmetric Effects without Collateral Constraints," *American Economic Journal: Macroeconomics*, 9, 222–66.

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

CHANEY, T., D. SRAER, AND D. THESMAR (2012): "The Collateral Channel: How Real Estate Shocks Affect Corporate Investment," *The American Economic Review*, 102, 2381–2409.

CHARI, V. V., P. J. KEHOE, AND E. R. MCGRATTAN (2007): "Business Cycle Accounting," *Econometrica*, 75, 781–836.

CHAVA, S. AND M. R. ROBERTS (2008): "How Does Financing Impact Investment? The Role of Debt Covenants," *The Journal of Finance*, 63, 2085–2121.

CHODOROW-REICH, G. AND A. FALATO (2017): "The Loan Covenant Channel: How Bank Health Transmits to the Real Economy," Working Paper 23879, NBER.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113, 1–45.

CHRISTIANO, L. J., M. S. EICHENBAUM, AND M. TRABANDT (2018): "On DSGE Models," Working Paper 24811, National Bureau of Economic Research.

Cloyne, J., C. Ferreira, M. Froemel, and P. Surico (2018): "Investment, Financial Frictions and the Dynamic Effects of Monetary Policy," Working paper.

Corbae, D. and E. Quintin (2015): "Leverage and the Foreclosure Crisis," *Journal of Political Economy*, 123, 1–65.

Cordoba, J.-C. and M. Ripoll (2004): "Credit Cycles Redux," *International Economic Review*, 45, 1011–1046.

Crouzet, N. (2017): "Aggregate implications of corporate debt choices," *The Review of Economic Studies*, 85, 1635–1682.

Cummins, J. G. and G. L. Violante (2002): "Investment-Specific Technical Change in the United States (1947–2000): Measurement and Macroeconomic Consequences," *Review of Economic Dynamics*, 5, 243 – 284.

Damodaran, A. (2012): *Investment valuation: Tools and techniques for determining the value of any asset*, vol. 666, John Wiley & Sons.

Del Negro, M., G. B. Eggertsson, A. Ferrero, and N. Kiyotaki (2016): "The great escape? A quantitative evaluation of the Fed's liquidity facilities," Working Paper 22259, National Bureau of Economic Research.

Den Haan, W. and T. Drechsel (2018): "Agnostic Structural Disturbances (ASDs): Detecting and Reducing Misspecification in Empirical Macroeconomic Models," Tech. rep., CEPR Discussion Papers.

Diamond, D. W., Y. Hu, and R. G. Rajan (2017): "Pledgeability, Industry Liquidity, and Financing Cycles," Tech. rep., National Bureau of Economic Research.

DiCecio, R. (2009): "Sticky wages and sectoral labor comovement," *Journal of Economic Dynamics and Control*, 33, 538 – 553.

Dichev, I. D. and D. J. Skinner (2002): "Large–Sample Evidence on the Debt Covenant Hypothesis," *Journal of Accounting Research*, 40, 1091–1123.

Falato, A. and N. Liang (2017): "Do Creditor Rights Increase Employment Risk? Evidence from Loan Covenants," *The Journal of Finance*, 71, 2545–2590.

Faust, J. (1998): "The robustness of identified VAR conclusions about money," *Carnegie-Rochester Conference Series on Public Policy*, 49, 207 – 244.

Faust, J. and E. M. Leeper (1997): "When Do Long-Run Identifying Restrictions Give Reliable Results?" *Journal of Business & Economic Statistics*, 15, 345–353.

Fisher, J. D. (2006): "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy*, 114, 413–451.

Francis, N., M. T. Owyang, J. E. Roush, and R. DiCecio (2014): "A Flexible Finite-Horizon Alternative to Long-Run Restrictions with an Application to Technology Shocks," *The Review of Economics and Statistics*, 96, 638–647.

Fuerst, T. S. (1995): "Monetary and Financial Interactions in the Business Cycle," *Journal of Money, Credit and Banking*, 27, 1321–1338.

Gali, J. (1996): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations," Working Paper 5721, National Bureau of Economic Research.

——— (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89, 249–271.

Galí, J. (2018): "The State of New Keynesian Economics: A Partial Assessment," Working Paper 24845, National Bureau of Economic Research.

Garleanu, N. and J. Zwiebel (2009): "Design and Renegotiation of Debt Covenants," *The Review of Financial Studies*, 22, 749–781.

Geanakoplos, J. (2010): "The Leverage Cycle," *NBER Macroeconomics Annual*, 24, 1–66.

Gertler, M. and S. Gilchrist (2018): "What Happened: Financial Factors in the Great Recession," *Journal of Economic Perspectives*, 32, 3–30.

Gertler, M. and P. Karadi (2011): "A model of unconventional monetary policy," *Journal of monetary Economics*, 58, 17–34.

Gomes, J., U. Jermann, and L. Schmid (2016): "Sticky Leverage," *American Economic Review*, 106, 3800–3828.

Greenwald, D. L. (2017): "The mortgage credit channel of macroeconomic transmission," *Working Paper*.

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

Greenwood, J., Z. Hercowitz, and P. Krusell (2000): "The role of investment-specific technological change in the business cycle," *European Economic Review*, 44, 91 – 115.

Grjebine, T., U. Szczerbowicz, and F. Tripier (2018): "Corporate debt structure and economic recoveries," *European Economic Review*, 101, 77 – 100.

Hart, O. and J. Moore (1994): "A Theory of Debt Based on the Inalienability of Human Capital," *The Quarterly Journal of Economics*, 109, 841.

Hayashi, F. (1982): "Tobin's Marginal q and Average q: A Neoclassical Interpretation," *Econometrica*, 50, 213–224.

Hennessy, C. A. and T. M. Whited (2005): "Debt dynamics," *The Journal of Finance*, 60, 1129–1165.

Iacoviello, M. (2005): "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle," *American Economic Review*, 95, 739–764.

Iacoviello, M. and S. Neri (2010): "Housing market spillovers: evidence from an estimated DSGE model," *American Economic Journal: Macroeconomics*, 2, 125–64.

Ingholt, M. M. (2018): "Multiple Credit Constraints and Time-Varying Macroeconomic Dynamics," *Unpublished Manuscript.*

Ivashina, V. and D. Scharfstein (2010): "Loan Syndication and Credit Cycles," *American Economic Review*, 100, 57–61.

Jaimovich, N. and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?" *American Economic Review*, 99, 1097–1118.

Jappelli, T. and M. Pagano (1989): "Consumption and capital market imperfections: An international comparison," *The American Economic Review*, 1088–1105.

Jensen, M. C. and W. H. Meckling (1976): "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of financial economics*, 3, 305–360.

Jermann, U. and V. Quadrini (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 102, 238–71.

Jordà, O. (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95, 161–182.

Jordà, O., M. Schularick, and A. M. Taylor (2017): "The effects of quasi-random monetary experiments," Working Paper 23074, National Bureau of Economic Research.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010): "Investment shocks and business cycles," *Journal of Monetary Economics*, 57, 132 – 145.

——— (2011): "Investment shocks and the relative price of investment," *Review of Economic Dynamics*, 14, 102 – 121, special issue: Sources of Business Cycles.

Kehoe, P. J., V. Midrigan, and E. Pastorino (2018): "Evolution of Modern Business Cycle Models: Accounting for the Great Recession," Working Paper 24741, National Bureau of Economic Research.

Kilian, L. and H. Lütkepohl (2017): *Structural vector autoregressive analysis*, Cambridge University Press.

Kiyotaki, N. (1998): "Credit and business cycles," *The Japanese Economic Review*, 49, 18–35.

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

——— (2012): "Liquidity, Business Cycles, and Monetary Policy," Working Paper 17934, National Bureau of Economic Research.

Kocherlakota, N. R. (2000): "Creating business cycles through credit constraints," *Quarterly Review*, 2–10.

Korinek, A. (2011): "Excessive dollar borrowing in emerging markets: Balance sheet effects and macroeconomic externalities," *Working paper*.

Li, H. (2016): "Leverage and Productivity," 2016 Meeting Papers 795, Society for Economic Dynamics.

Lian, C. and Y. Ma (2018): "Anatomy of Corporate Borrowing Constraints," *Working paper*.

Lintner, J. (1956): "Distribution of incomes of corporations among dividends, retained earnings, and taxes," *The American Economic Review*, 46, 97–113.

Liu, Z., P. Wang, and T. Zha (2013): "Land-Price Dynamics And Macroeconomic Fluctuations," *Econometrica*, 81, 1147–1184.

Lorenzoni, G. (2008): "Inefficient Credit Booms," *The Review of Economic Studies*, 75, 809–833.

Mendoza, E. G. (2006): "Lessons from the Debt-Deflation Theory of Sudden Stops," *The American Economic Review*, 96, 411–416.

Murfin, J. (2012): "The Supply-Side Determinants of Loan Contract Strictness," *The Journal of Finance*, 67, 1565–1601.

Nini, G., D. C. Smith, and A. Sufi (2012): "Creditor Control Rights, Corporate Governance, and Firm Value," *The Review of Financial Studies*, 25, 1713–1761.

Pfeifer, J. (2016): "Macroeconomic Effects of Financial Shocks: Comment," Dynare Working Papers 50, CEPREMAP.

Quadrini, V. (2011): "Financial frictions in macroeconomic fluctuations," *Economic Quarterly, Federal Reserve Bank of Richmond*, 97, 209–254.

Ramey, V. (2016): "Chapter 2 - Macroeconomic Shocks and Their Propagation," Elsevier, vol. 2 of *Handbook of Macroeconomics*, 71 – 162.

Riddick, L. A. and T. M. Whited (2009): "The Corporate Propensity to Save," *The Journal of Finance*, 64, 1729–1766.

Roberts, M. R. and A. Sufi (2009a): "Control rights and capital structure: An empirical investigation," *The Journal of Finance*, 64, 1657–1695.

——— (2009b): "Renegotiation of financial contracts: Evidence from private credit agreements," *Journal of Financial Economics*, 93, 159 – 184.

Schmitt-Grohe, S. and M. Uribe (2012): "What's news in business cycles?" *Econometrica*, 80, 2733–2764.

——— (2016): "Adjustment to Small, Large, and Sunspot Shocks in Open Economies With Stock Collateral Constraints," Working Paper 22971, National Bureau of Economic Research.

Shleifer, A. and R. W. Vishny (1992): "Liquidation Values and Debt Capacity: A Market Equilibrium Approach," *The Journal of Finance*, 47, 1343–1366.

Smets, F. and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

Standard & Poor's Global Ratings (2013a): "General: Corporate Methodology," *Available at https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/8314109*.

——— (2013b): "General: Corporate Methodology: Ratios And Adjustments," *Available at https://www.standardandpoors.com/en_US/web/guest/article/-/view/sourceId/8330212*.

Strebulaev, I. A. and T. M. Whited (2012): "Dynamic Models and Structural Estimation in Corporate Finance," *Foundations and Trends in Finance*, 6, 1–163.

Sufi, A. (2009): "Bank Lines of Credit in Corporate Finance: An Empirical Analysis," *Review of Financial Studies*, 22, 1057–1088.

Townsend, R. M. (1979): "Optimal contracts and competitive markets with costly state verification," *Journal of Economic Theory*, 21, 265 – 293.

Uhlig, H. (2004): "Do technology shocks lead to a fall in total hours worked?" *Journal of the European Economic Association*, 2, 361–371.

Xiao, J. (2018): "Capital Allocation and Investment Dynamics in Credit Crises," *Unpublished Manuscript.*

Chapter 2

Commodity Booms and Busts in Emerging Economies

2.1 Introduction

Emerging economies, particularly those that are dependent on commodity exports, have a long history of volatile and disruptive economic cycles. A rich literature in International Macroeconomics has proposed several explanations for these cycles, pointing to different plausible triggers or underlying sources of shocks. The relative importance of the various triggers, however, still divides the literature. Aguiar and Gopinath (2007) argue that the main source of fluctuations is nonstationary total factor productivity (TFP) shocks - the cycle is the trend. García-Cicco et al. (2010) refute the argument, showing that these shocks only explain a negligible fraction of fluctuations. They contend that the main drivers of shocks are stationary TFP shocks as well as exogenous shocks to the country's interest rate. The latter result relates to work by Guimaraes (2011), Neumeyer and Perri (2005), and Uribe and Yue (2006), who highlight the role of changes in global interest rates as a potential driver of the cycle. The role of commodity prices and, more generally, terms of trade, has been equally divisive. Mendoza (1995) and Kose (2002) argue that fluctuations in the terms of trade explain a large fraction of the output variance. However, empirical work by Schmitt-Grohe and Uribe (2017) has raised questions on the ability of terms of trade to match critical features of business cycles in emerging economies. Interestingly, though, estimates by Fernández et al. (2017) suggest that fluctuations in commodity prices account for a significant fraction of output fluctuations.[1] For

[1]Schmitt-Grohe and Uribe (2017) empirically estimate the impulse response functions of GDP and consumption to terms of trade shocks. They find that consumption responds negatively to terms of trade innovations, in sharp contrast to the positive response of GDP. Given the overall positive comovement between consumption and GDP in the data, their work bodes negative prospects for terms of trade as a key driver of the cycle. Empirical results in Fernández et al. (2017) however, suggest that commodity prices potentially account for a significant fraction of output fluctuations, though their paper does not provide impulse response functions for the various macroeconomic aggregates to shed light on the comovements across variables and potential mechanisms. Another empirical paper with a focus on commodity prices, and the resulting procyclicality of fiscal policy, is Cespedes and Velasco (2014).

economies with a comparative advantage in the production of commodities, fluctuations in the terms of trade and in real commodity prices tend to display a highly positive correlation, and hence the tension between these two empirical studies' results invites a fresh take. In turn, these results call for a tighter connection with the aforementioned studies on the relative importance of different productivity and interest rate shocks.

This paper seeks to quantitatively assess the drivers of emerging economy business cycles using a unified model that nests the various sources of shocks advanced in the literature. The model builds on the small open economy setting of Aguiar and Gopinath (2007) and García-Cicco et al. (2010) by adding two elements absent from their analysis. First, it allows for a second sector to capture the separate role of commodities in the economy. Specifically, the analysis focuses on the case of a net commodity exporting country facing exogenous international price changes. Second, the model embeds a negative relation between the interest rate premium and commodity prices. The relevance of this channel has recently been highlighted by Fernández et al. (2015) and Shousha (2016), and is consistent with the empirical evidence.

To study the predictions of our model, we resort both to a calibration exercise and to the estimation of the model with Bayesian methods. The quantitative analysis throughout the paper focuses on Argentina, a quintessential example of a commodity exporting emerging economy. Given the lengthy duration of Argentine cycles, we carry the analysis over a long period (1900-2015) in order to capture multiple cycles.[2] To set the stage, we begin by revisiting a number of empirical regularities. In common with other emerging economies, Argentina displays large and persistent cyclical fluctuations, excess volatility of consumption over output, high volatility of investment, and a negative correlation between output growth and the trade balance. In addition, the Argentine data reveal large positive effects of world commodity price shocks on output, consumption, and investment, as well as negative effects on the trade balance. We identify these shocks using a structural vector autoregression (SVAR) model with a standard Cholesky decomposition, relying on the assumption that world commodity prices are not contemporaneously affected by Argentina's economic activity. Furthermore, the data display a strong negative association between interest rate spreads in Argentina and world commodity prices. Maintaining the assumption that international commodity prices are exogenous to developments in Argentina's economy, we estimate this relation with a set of regressions of measures of Argentine real rates (net of world interest rates) on an international commodity price index and various controls. The strongly negative relation is robust across a

[2]Shousha (2016) focuses on a quarterly sample from 1994-2013 pooling together various emerging economies. In the case of Argentina, this would not be lengthy enough to capture a full cycle. Aguiar and Gopinath (2007) analyze an even shorter period for Argentina, 1993-2002. Fernández et al. (2015) estimate their model on a pool of countries (Brazil, Chile, Colombia, and Peru) covering the period 2000:Q1 to 2014:Q3. We concur with García-Cicco et al. (2010) in that a long period is necessary in order to distinguish trend and cyclical shocks. They base the analysis on 1900-2010 and hence our results are more directly comparable to theirs.

number of specifications, with different spread measures and different sets of controls, including output growth, the trade balance and the debt-to-GDP ratio. The lower bound of our estimates suggests that a 10% deviation of commodity prices from their long-run mean can move Argentina's real interest rate spread by almost 2 percentage points. This finding confirms the existing evidence from the literature on interest rate spreads of commodity exporting economies (see in particular Bastourre et al., 2012, Fernández et al., 2015, and Shousha, 2016). It also connects with earlier work by Kaminsky et al. (2005) on the procyclicality of capital flows in developing countries.[3]

In the model calibration exercise we analyze the response of the economy to commodity price shocks of a sensibly calibrated size, which we can directly compare to the impulse response functions obtained from the SVAR. We find that the model impulse response functions line up well with their empirical counterparts. The two effects stemming from commodity prices (that is, the competitiveness effect and the borrowing cost effect) *jointly* produce impulse response functions to a commodity price shock that match the empirical responses. They generate strongly positive effects on GDP, consumption, and investment, and a negative effect on the total trade balance. They also give rise to a somewhat larger response of consumption over output. We show that the first effect alone (akin to a productivity increase) cannot generate a countercyclical trade balance. Similarly, the second effect alone (which is isomorphic to a simple negative interest rate shock) does not give a contemporaneous response in output, while consumption and investment do increase on impact. The net contribution of the two effects can reproduce the empirical regularities.

The aim of the structural estimation of the model is to gauge the quantitative importance of commodity price shocks, relative to other shocks, in driving the business cycle. We apply Bayesian estimation methods, using data on output, consumption, investment, and the trade balance of Argentina. We estimate the stochastic processes of various exogenous disturbances, as well as the two parameters governing the sensitivity of the interest rate spread to commodity prices and to the debt level. Our results suggest a sizeable contribution of commodity price shocks to Argentine business cycle fluctuations. The posterior forecast error variance decomposition based on data from 1900 to 2015 attributes 22% of the observed variation in output growth to commodity price shocks. Furthermore, 24% of consumption growth and 34% of investment growth can be accounted for by commodity price shocks. Reassuringly, the model-implied process for the commodity price shares important features with empirically observed world commodity prices. Since it mimics the data particularly closely after 1950, we carry out the estimation on the post-1950 subsample and find that the contribution of commodity price shocks to output, consumption, and investment growth rises to around 38%, 42%, and 61%, respectively.

Our assessment of the remaining variation in macroeconomic aggregates sheds additional light on the debate about the candidate drivers of emerging economy business

[3]See also Reinhart and Reinhart (2009), Gavin et al. (1996), Prasad et al. (2006), and Frankel (2011).

cycles previously proposed in the literature. We find that, in general, stationary technology shocks remain the most important source of fluctuations, explaining around half of the variation in output growth. These stationary shocks to TFP are quantitatively more important than non-stationary TFP shocks. While this echoes the conclusion of García-Cicco et al. (2010), who question the notion that the "cycle is the trend" in emerging economies, the contribution of nonstationary shocks remains non-negligible, as these shocks are able to explain 21% of the variation in output growth in both samples used in the estimation.[4] We also find a significant role for preference shocks and interest rate shocks in explaining the variation in consumption, investment, and the trade balance.

Taken together, our results suggest that commodity prices should feature prominently in the analysis of business cycles in emerging economies. In terms of quantitative contribution, they are among the three most important shocks driving output growth in Argentina. Importantly, shocks to international commodity prices, in contrast to inherently more opaque concepts such as domestic TFP shocks, are factors that are easier to identify and measure, and potentially act upon, by policy makers.[5]

The rest of the paper is organized as follows. Section 2.2 presents a number of empirical regularities characterizing Argentine business cycles. As said, many of these regularities are shared with other emerging commodity exporting countries, though for the sake of accuracy in the mapping from the data to the model, we think it is insightful to focus on a single country. Section 2.3 introduces the model. Section 2.4 performs the calibration exercise and studies the role of commodity price shocks in the model. Section 2.5 estimates the model and carries out a quantitative analysis of the various sources of shocks; it also discusses practical issues concerning the measurement of real GDP. Section 2.6 contains concluding remarks.

2.2 Emerging Market Cycles: Empirical Regularities

This section presents the main empirical features that characterize the business cycle of Argentina's economy from 1900 to 2015.

2.2.1 Data and Sample

Although there are strong commonalities across emerging countries, we think it is important to work with a straight mapping from a single country to the model, rather than using averages across different countries, which might confound effects due to aggregation. The focus on a long time period is both insightful and befitting

[4]Our conclusion with respect to this aspect is quite similar to recent findings of Akinci (2017).

[5]Our model does not feature sovereign default or distress. While sovereign default episodes have been important for Argentina, we think there is a lot of merit in understanding the triggers of the cycles and how they are affected by external factors such as commodity prices in a relatively simple setting, which more realistically would end with a technical default. A better understanding of these regularities may actually help in avoiding default episodes by guiding policy. As will become clear, the model features a negative externality, as households do not take into account the effect of their borrowing on interest rates, which can lead to overborrowing.

for a number of reasons. First, Argentina's large and persistent economic cycles call for a lengthy time span in order to capture a reasonable number of completed cycles in the analysis. Second, unlike advanced economies, Argentina's cyclical properties have shown virtually no changes over this long period. This is apparent in Figure 2.1, Panel (a), which plots the logarithm of Argentine real GDP per capita from 1900 to 2015. Argentina's output volatility in the first half of the 20th century (measured as the standard deviation of real GDP growth rates) is practically the same as the volatility in the post 1950 period, despite the higher levels of development in the latter part of the sample. In the corresponding plot for the United States, shown in Panel (b), marked changes in the volatility of output are visible. This typically leads researchers to separately analyze data before and after the World War II, or before and after the 1980s, which was when the Great Moderation occurred in the United States. Such changes in volatility are not present in Argentina, which supports the case for analyzing fluctuations jointly over the entire period.[6] Third, Argentina's trend growth rate has been remarkably stable since 1900, at 1.2% per year, a constancy that can be fully appreciated by taking a long-term perspective in analyzing its business cycles.[7] In addition to output data, we will focus on typical macroeconomic variables of interest in small open economies, by studying the fluctuations of consumption, investment, and the trade balance. The data come from a variety of sources, including most notably Ferreres (2005).[8]

Furthermore, since our aim is to assess the importance of commodity price fluctuations for Argentina's economy, we need to select an appropriate commodity price index. Our preferred index is the one constructed by Grilli and Yang (1988), which we update following Pfaffenzeller et al. (2007). The index is available from 1900 and reflects *world* commodity prices, which is advantageous because developments in global prices are arguably exogenous to economic conditions in Argentina (see further discussion below). The drawback, of course, is that it may capture price developments of commodities that are unimportant, or even absent, in Argentina's commodity export composition.[9] We therefore cross-check this index with an *Argentina-specific* commodity price index, which we construct using commodity price data provided by the World Bank, together with trade weights available from the UN Comtrade data base. This construction is possible from 1962 onwards. Figure 2.2, Panel (a), plots the two indices (in nominal terms) and shows that their year-on-year changes

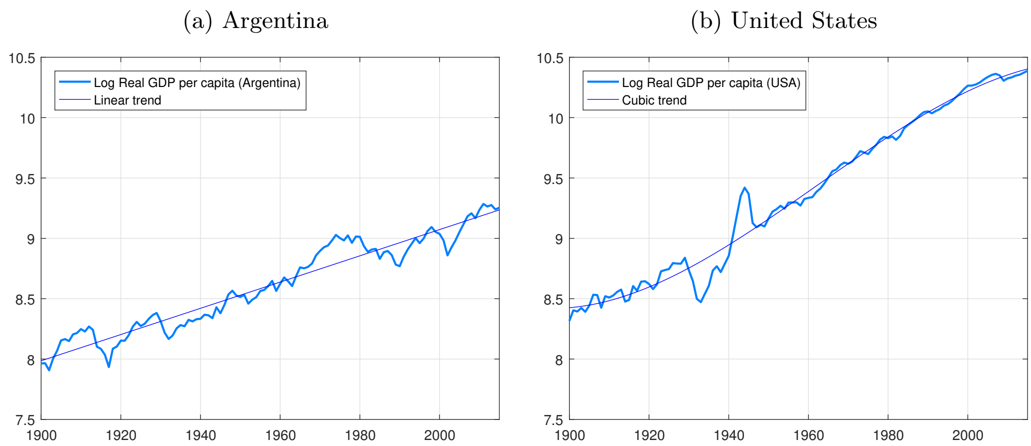
[6] A similar argument is made by García-Cicco et al. (2010); they emphasize the importance of a long horizon to disentangle transitory shocks from shocks to trend growth in business cycles of emerging economies, which are the focus of Aguiar and Gopinath (2007). We will also aim at disentangling these two types of shocks in our model estimation, in addition to our focus on commodity prices.

[7] This is also different in the US, where low frequency changes in the trend growth rate are present (see Antolin-Diaz et al., 2017, for comprehensive evidence). We therefore fit a cubic rather than linear trend in Panel (b) of Figure 2.1.

[8] We extend the series of Ferreres (2005) to 2015. Compared to García-Cicco et al. (2010), we add another half decade of data. Details on the sources and construction of the data are provided in Appendix 2.7.1.

[9] Argentina exports mainly agricultural and food commodities such as meat, maize, and soy beans, but to a lesser extent also petroleum, gold, and other non-food commodities.

Figure 2.1: OUTPUT PER CAPITA 1900-2015 - ARGENTINA VS. US



Note: Panel (a) displays Argentine real GDP per capita in log scale, together with a linear trend. Panel (b) shows the log of US real GDP per capita and adds a cubic trend. The sources are Ferreres (2005) (updated series) and the US Bureau of Economic Analysis, respectively.

are fairly synchronized, mitigating the concern that the world price index may not be representative of commodity prices faced by Argentina. We deflate the Grilli and Yang (1988) index to be a relative ("real") price using an index of (US-dollar denominated) import prices for Argentina.[10] Figure 2.2, Panel (b) plots this time series in deviations from its sample mean. We focus on mean deviations rather than other detrending methods, since we are interested in capturing persistent movements over longer time spans, sometimes referred to as "supercycles" in commodity prices.

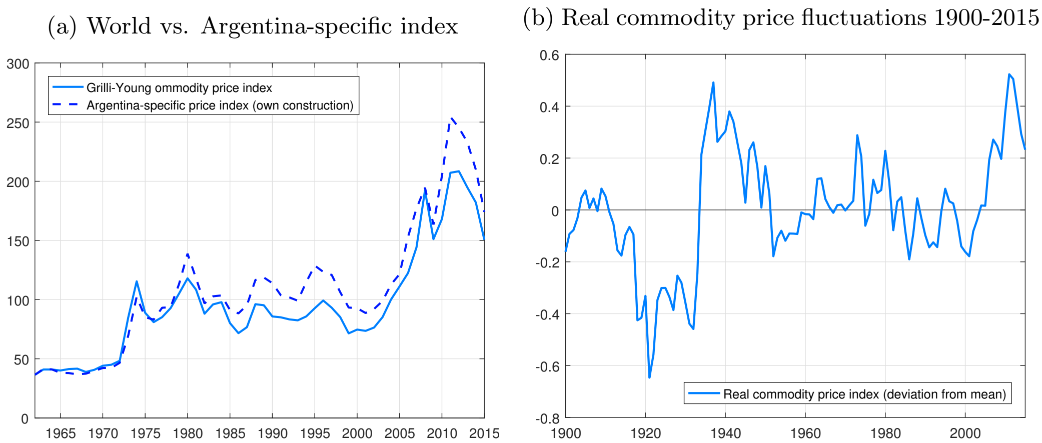
We begin our characterization of the empirical regularities by documenting business cycle moments. We then turn to estimating an SVAR in order to gauge the dynamic effects of exogenous commodity price developments on Argentina's economy. Furthermore, we present evidence on the relation of commodity prices and Argentina's real interest rate spread. Finally, we summarize the insights of this section into a set of stylized facts.

2.2.2 Business Cycle Moments

Table 2.1 summarizes key business cycle moments of Argentina's economy. We report mean, standard deviation, persistence, and contemporaneous cross-correlation of GDP growth, consumption growth, investment growth (all per capita), as well as the trade balance, defined as exports minus imports scaled by GDP. As the table shows, many properties of the Argentine business cycle are in line with what is typically observed in advanced economies. Output, consumption, and investment are strongly correlated and investment is much more volatile than output. On the other

[10] The import price index updates the series published by Ferreres (2005). We have tried alternative ways of deflating the commodity price series, for example using manufacturing prices (also expressed in US dollars), or the US consumer price index. The changes did not have a material impact on the results we present. We prefer the deflation using import prices (expressed in US dollars), since this brings the observed price index closest to the corresponding concept in our model, which is the relative price between commodities and a final tradable consumption good.

Figure 2.2: Commodity prices



Note: Panel (a) compares the updated index of Grilli and Yang (1988) with an Argentina-specific commodity price index constructed based on UN Comtrade and World Bank data. These series are in nominal terms and normalized to the same value in 1962. Panel (b) displays the commodity price index of Grilli and Yang (1988), deflated with the Argentine import price index (in US dollars), and in log-deviations from its sample mean.

hand, there are features that are distinctive of fluctuations in emerging markets. In particular, it is worth highlighting that consumption growth is more volatile than output growth.[11] Furthermore, as often observed in emerging markets, the trade balance is countercyclical. In the case of Argentina the contemporaneous correlation with output growth is not large, calculated at -0.07, but the magnitude of the negative correlation is more pronounced with consumption and investment.

Table 2.1: Business Cycle Moments 1900-2015

	GDP growth	**Cons. growth**	**Inv. growth**	**Trade balance**
Mean	1.17%	1.12%	1.40%	-0.04%
Standard deviation	5.27%	5.84%	19.16%	4.76%
Persistence	0.14	0.05	0.34	0.72
Correlation with GDP growth	1	0.86	0.76	-0.07
Correlation with Cons. growth	0.86	1	0.49	-0.11
Correlation with Inv. growth	0.76	0.49	1	-0.20
Correlation with trade balance	-0.07	-0.11	-0.20	1

Note: GDP, consumption, and investment growth are real and in per capita terms. The trade balance is defined as total exports minus total imports, scaled by GDP. Persistence is the coefficient from an estimated AR(1) process. The frequency of the data is annual.

[11] Interestingly, the excess volatility of consumption is smaller in our sample than in García-Cicco et al. (2010)'s sample, suggesting that this phenomenon has attenuated in recent years.

2.2.3 Commodity Price Shocks and Emerging Economy Cycles

In order to gauge the effect of international commodity prices on merging market business cycles, we consider the following structural vector autoregression (SVAR):

$$A_0 Z_t = at + A_1 Z_{t-1} + \ldots + A_p Z_{t-p} + u_t, \tag{2.1}$$

where Z_t is a vector containing the commodity price index in log deviations from mean, as plotted in Figure 2.2, together with the log-levels of the business cycle variables of interest - output, consumption, investment, and the trade balance; u_t is a vector of normally distributed structural shocks with covariance matrix $\mathbb{E}(u_t u_t') = I_5$; and t is a linear time trend. We set the number of lags to $p = 2$.[12]

We estimate the reduced form version of equation (2.1) using OLS, obtain the residuals $\hat{\epsilon}_t = \hat{A}_0^{-1}\hat{u}_t$ and then recover commodity price shocks, that is, the element of $\hat{u}_t$ corresponding to commodity prices, using restrictions on A_0. Our underlying identifying assumption is that international commodity prices are not contemporaneously affected by any other variable in the system. Given that Argentina is a relatively small country that should not be a driver of world-wide commodity prices, we believe this assumption is reasonable and justifies ordering the commodity price first in a Cholesky decomposition of the covariance matrix of u_t.[13]

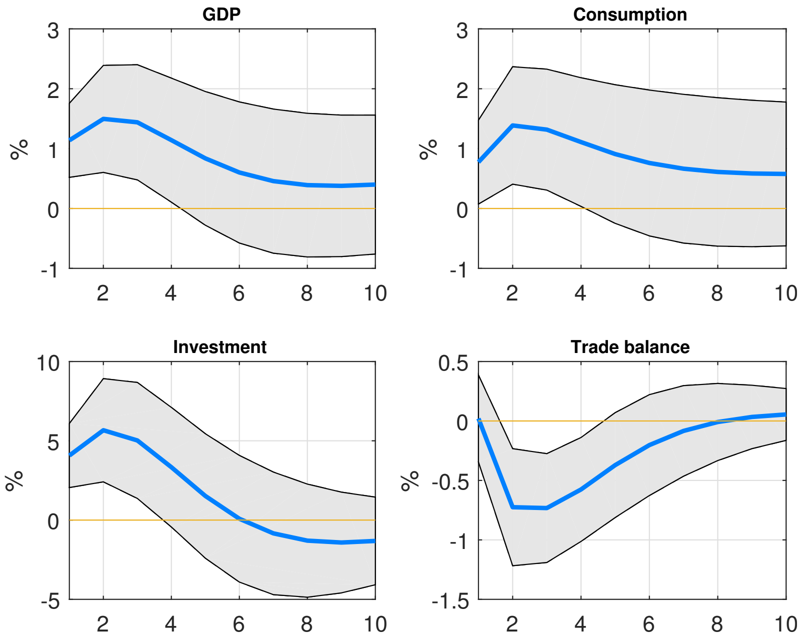
Due to the imperfections in the measurement of commodity prices faced by Argentina discussed earlier, we focus solely on the IRFs of the SVAR, but do not resort to a forecast error variance decomposition. Our working assumption is that the IRFs in response to a shock identified from this specification give a meaningful representation of the dynamics following an exogenous shock to international commodity prices. However, we think that making quantitative statements about the total contribution of commodity prices to the variance of output from this exercise could be misleading given the noisy nature of the Grilli and Yang (1988) index as a measure of the actual price movements faced by Argentina. We instead carry out such a decomposition using the structural model in Section 2.5.

The impulse response functions to a one standard deviation shock to commodity prices are plotted in Figure 2.3. The results show that there is a statistically and economically significant positive response of output, consumption, and investment. The total trade balance response is negative, that is, net exports fall in response to a commodity price increase. All responses are hump-shaped, peaking around two years following the shock, and quite persistent. Measured at peak, a one standard deviation shock in international commodity prices increases the level of real GDP per capita by more than one percent.

[12] This lag length is selected against $p = 1$ using various lag length selection criteria.

[13] We leave the remaining shocks to the system unidentified, so that the ordering of the remaining variables is irrelevant.

Figure 2.3: Impulse responses to 1 S.D. Commodity Price Shock



Note: The structural shock is identified using Cholesky ordering. 80% confidence bands are plotted, as suggested by Sims and Zha (1999). GDP, consumption, and investment are real, in per-capita terms and in log-levels. The trade balance is defined as exports net of imports divided by GDP.

2.2.4 Commodity Prices and Interest Rate Spreads

What are possible channels behind the influence of commodity prices on emerging market business cycles? One key observation that has been highlighted in previous research on commodity exporting economies is the strong negative comovement of interest rate spreads and commodity prices. Fernández et al. (2015) highlight the strong negative effect of commodity price increases on country risk premia in sovereign bond spreads. Bastourre et al. (2012) estimate the correlation between a common factor of emerging economy bond returns and a common factor of commodity prices to be -0.81. Shousha (2016) emphasizes that the negative correlation is a major difference between emerging and advanced commodity exporters. Incorporating this effect into our analysis is important, since strongly countercyclical interest rate movements in general have been found to be a key driver of emerging markets business cycles, see for example Uribe and Yue (2006) and Neumeyer and Perri (2005).[14]

To shed further light on the link between the real spread and commodity prices in the case of Argentina, we run a set of regressions of the Argentine real interest rate spread on the real commodity price index (in log deviations from its mean). The regressions are specified as follows:

[14]This result connects with work on the procyclicality of capital flows and borrowing in emerging and developing economies. See for example Kaminsky et al. (2005).

$$r_t - r_t^* = \alpha + \xi(ln\tilde{p}_t - ln\bar{\tilde{p}}) + \beta X_t + v_t, \tag{2.2}$$

where r_t is the real interest rate of Argentina, r_t^* is a measure of the world interest rate, $\tilde{p}_t$ is the commodity price (with $ln\tilde{p}_t - ln\bar{\tilde{p}}$ being the log deviation from mean, which we plot in Figure 2.2, Panel (b)), and X_t is a vector of control variables including output growth, the debt-to-GDP ratio and the trade balance. The key parameter of interest is ξ, which denotes the sensitivity of the real interest rate spread with respect to changes in world commodity prices. Note that this sensitivity parameter will also feature in our model and we will calibrate it based on the results presented in this section. Since interest rate data for Argentina are not available over our baseline 1900-2015 sample, we stick to a smaller time period and try different interest rate series available. Specifically, we use the domestic lending rate, savings rate, and the money market rate, which are all provided by the IMF International Financial Statistics in nominal terms. To obtain a real measure we deflate these series using a corrected inflation measure for Argentine inflation ("inflación verdadera"), since several authors have highlighted the misreporting of inflation by official sources in recent years (see Cavallo, 2013, for a discussion).[15] For the world interest rate we use a measure of the UK real interest rate published by the Bank of England. We once again emphasize that the commodity price measure captures international commodity price developments which are arguably exogenous to economic activity in Argentina.

The baseline results are presented in Table 2.2. We show several other results using different interest rate measures in Appendix 2.7.2. Our findings across all regressions, including those in the appendix, give negative point estimates of ξ. These estimates are economically significant though not always statistically significant, likely due to the small sample. If we consider the smallest estimate (in absolute value) that is statistically significant, which is -0.199, the interpretation is that a 10% deviation of commodity prices from their long-run mean can move Argentina's real interest spread by almost 2 percentage points. We view this as strong evidence in support of a channel by which exogenous international commodity prices put downward pressure on interest rate premia faced by commodity exporting emerging economies. This evidence will guide our modeling choices below, where we also provide further theoretical discussion of this economic relation.

2.2.5 Summary of Stylized Facts

Based on the empirical analysis above, we summarize the following stylized facts around aggregate fluctuations in Argentina 1900-2015:

1. A relatively constant trend in GDP per capita growth at an average of 1.2%

[15] In a previous version of the paper we additionally used a real interest rate measure directly provided by the world bank. This series is also based on the IMF lending rate measure but uses the official Argentine GDP deflator to obtain a real series, which we chose to avoid. The results are available on request.

Table 2.2: Regression Results

	(1)	(2)	(3)	(4)	(5)
LHS variable	Real spread (calculated from domestic lending rate)				
Commodity price	-0.200***	-0.199***	-0.214***	-0.210***	-0.203***
	(0.049)	(0.045)	(0.051)	(0.051)	(0.050)
Output growth		-0.434**			-0.406
		(0.206)			(0.241)
Trade balance			-0.252		-0.164
			(0.224)		(0.385)
Debt-to-GDP ratio				-0.033	0.015
				(0.036)	(0.062)
Constant	0.023*	0.034**	0.024*	0.041*	0.026
	(0.012)	(0.012)	(0.012)	(0.024)	(0.034)
Observations	21	21	21	21	21
R-squared	0.462	0.568	0.497	0.485	0.573

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The real spread is calculated by deflating the domestic lending rate, provided by the IMF, with a corrected inflation measure (see Cavallo, 2013), and then subtracting the UK real rate. The commodity price is in log deviations from mean, as plotted in Figure 2.2, Panel (b). Appendix 2.7.1 provides details on the sources of the other regressors.

annually, with a relatively stable variance throughout the period.

2. Excess volatility of consumption over output.
3. A negative correlation between GDP growth and the trade balance.
4. Large of effects of commodity price shocks on all key business cycle variables.
5. A negative relation between interest spreads and commodity prices.

2.3 A Two-Sector Small Open Economy Model

We build on the small open economy model formulated by Aguiar and Gopinath (2007) and García-Cicco et al. (2010), which in turn build on Mendoza (1991).[16] Our model adds two elements absent in their analysis. First it allows for a second sector to capture the distinctive role of commodities present in many emerging economies. Second, as in Shousha (2016), the model embeds a negative relation between the interest rate premium and commodity prices, consistent with the empirical evidence presented above. The model nests the various sources of shocks identified in previous work and allows for a double-role of commodity prices. Increases in commodity prices improve both the competitiveness of the economy (which is a net commodity

[16]We abstract from nominal frictions and the important question of fixed versus nominal exchange rate choice. See for example Frankel (2004), and Mitchener and Pina (2016), who examine the costs and benefits of fixed exchange rates. For a modeling framework that incorporates nominal elements, we refer readers to Gali and Monacelli (2005) and the literature that built on their seminal contribution.

exporter) and the economy's borrowing terms, as higher prices are associated with lower spreads between the country's borrowing rates and world interest rates.

We begin by describing the technology. There are two sectors in the economy: a final-good sector and a commodity-producing sector. The final good is produced by combining capital K_t^1, commodity inputs $\tilde{M}_t$, and labor N_t^1. It can be consumed, invested and exported or imported. The production function in the final good sector is

$$Y_t = a_t (K_t^1)^{\alpha_K} (\tilde{M}_t)^{\alpha_M} (X_t N_t^1)^{1-\alpha_K-\alpha_M}. \tag{2.3}$$

Commodities can be produced domestically using capital K_t^2 and labor N_t^2; they can be used as an intermediate input in final goods production or traded on international markets. The production function in the commodity sector is

$$\tilde{Y}_t = \tilde{a}_t (K_t^2)^{\tilde{\alpha_K}} (X_t N_t^2)^{1-\tilde{\alpha_K}}. \tag{2.4}$$

In the production functions, a_t and $\tilde{a}_t$ capture total factor productivities, which are exogenous and assumed to be stationary. X_t is the nonstationary level of labor-augmenting technology common to both sectors. We denote the gross growth rate of the nonstationary technology as $g_t = X_t/X_{t-1}$, which is stochastic with mean g. X_t is introduced to capture shocks to the trend, which has been a key focus in the literature on emerging market business cycles.[17] The price of the final good is normalized to 1 and the price of commodities $\tilde{p}_t$ is exogenously given on world markets and subject to shocks. We assume that a_t, $\tilde{a}_t$, g_t, and $\tilde{p}_t$ follow stochastic processes which will be specified further below.

Firms in both sectors rent capital and hire labor in competitive input markets. The total stock of capital in the economy K_t is measured in final goods and is divided between the two production technologies, so that

$$K_t = K_t^1 + K_t^2. \tag{2.5}$$

Capital depreciates at rate δ and is accumulated through investment I_t which gives

$$K_{t+1} = (1-\delta)K_t + I_t. \tag{2.6}$$

The economy is populated by a representative household who supplies the two types of labor, owns and rents out the capital stock, and borrows from abroad. The budget constraint is given by

$$C_t + K_{t+1} + D_t + S_t + \frac{\phi}{2}\left(\frac{K_{t+1}}{K_t} - g\right)^2 = r_t^{k1}K_t^1 + r_t^{k2}K_t^2 + w_t^1 N_t^1 + w_t^2 N_t^2 + (1-\delta)K_t + \frac{D_{t+1}}{1+r_t}, \tag{2.7}$$

[17]See in particular Aguiar and Gopinath (2007). The fact that in our model the nonstationary technology is common to both sectors ensures that the model admits a non-stochastic balanced growth path (BGP), as shown in Appendix 2.7.3.

where C_t is final good consumption, D_t denotes the level of (real) debt, and $\frac{D_{t+1}}{1+r_t}$ is newly issued debt at net interest rate r_t. S_t is exogenous government spending, where $s_t = S_t/X_{t-1}$ will follow a stochastic process to be specified further below. r_t^{kj} and w_t^j, $j = 1, 2$, are the returns from renting out capital and supplying labor to the two sectors, respectively. Note that in equilibrium the expected return on capital will equalize across the two sectors. The presence of $\phi > 0$ captures investment adjustment costs faced by the household.

The household's objective is to maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \nu_t \beta^t \frac{[C_t - \theta\omega^{-1}X_{t-1}(N_t^1)^{\omega} - \theta\tilde{\omega}^{-1}X_{t-1}(N_t^2)^{\tilde{\omega}}]^{1-\gamma} - 1}{1-\gamma} \tag{2.8}$$

with $\gamma > 0$, subject to the relevant constraints and a no-Ponzi condition. The parameter β is the discount factor and ν_t captures shocks to preferences. The utility function features Greenwood et al. (1988) preferences, which eliminate the wealth effect on labor supply. Note that the presence of X_{t-1} ensures a constant labor supply along the non-stochastic BGP. The Frisch elasticity of labor supply will be determined by ω and $\tilde{\omega}$, and θ governs the weight on the relative disutility of labor.

Based on the small open economy assumption, the steady state real interest rate is exogenously given. In particular, r_t is determined by the world interest rate r^* and a spread (or premium) term which is further composed of three additive terms:

$$r_t = r^* + \psi\big(e^{D_{t+1}^*/X_t - d^*} - 1\big) + \xi\left(ln(\tilde{p}_t) - ln(\tilde{p})\right) + \big(e^{\mu_t - 1} - 1\big)\,. \tag{2.9}$$

The first term of the spread in (2.9) is standard in the literature. Following Schmitt-Grohe and Uribe (2003), it is assumed that the premium is increasing in the (detrended) level of debt. The presence of D_{t+1}^* is taken as exogenous by the representative household but $D_{t+1} = D_{t+1}^*$ holds in equilibrium. This debt-elastic interest rate ensures a stationary solution of the model after detrending.[18]

The second term determining the spread $r_t - r^*$ captures the robust empirical observation, discussed in detail in Section 2.2.4, that commodity prices strongly affect interest rate premia of commodity exporting economies. The parameter ξ governs the sensitivity of the interest rate spread with respect to commodity price deviations from steady state and can be calibrated to the corresponding parameter we estimated in Section 2.2.4. Our approach here is to embed the relation between $r_t - r^*$ and $\tilde{p}_t$ in a reduced-form fashion, similar to Shousha (2016) and Fernández et al. (2015), who also document further empirical evidence in line with our findings. While we do not provide a complete formal rationalization of the relationship and focus mainly on the resulting implications for emerging economy business cycles, the link between commodity prices and interest rate premia can be derived from first principles following different approaches. Specifically, the negative relation between $r_t - r^*$ and $\tilde{p}_t$ may result from the effect of commodity prices on the country's repayment capacity to international creditors. This could come in the form of a borrowing constraint,

[18] See also Lubik (2007) for further discussion.

in which the value of the country's collateral depends directly on commodity prices through export earnings. Creditors decrease the required interest rate premium when commodity prices increase, as the collateral value of the economy is higher.[19] Min et al. (2003) provide empirical evidence for this particular channel, showing that export earnings and better repayment capacity bring down yield spreads. Alternatively, a possible mechanism could entail financial frictions in which domestic firms (rather than the government) borrow against collateral, which is positively linked to the terms of trade, and a relaxation in these constraints leads to a fall in credit spreads.[20]

Finally, the last term in the rate spread in (2.9) allows for a simple interest rate premium shock, similar to the one specified in García-Cicco et al. (2010). Since it is central to our objective to trace out the effects of commodity price movements for the economy, we also allow for the presence of μ_t in order to capture possible exogenously driven movements in the interest premium that are unrelated to commodity prices and thereby avoid hardwiring into the model that interest rate movements must be related to commodity prices. An alternative interpretation of this shock is of course an innovation in global interest rates (rather than the interest rate premium). We do not take a strong stance on this distinction in the analysis. From the domestic economy's perspective, exogenous changes in the premium and the global interest rate have similar effects on the domestic interest rate.

Our modeling choice is arguably restrictive, as apart from commodity prices we only allow one additional shock to directly affect interest rates via the last term in the spread. This restrictiveness has the benefit of allowing a direct comparison of the relative importance of the mechanism we introduce vis-à-vis a collection of exogenous disturbances which are defined in the same way as in García-Cicco et al. (2010). These authors also estimate their model on Argentine data over a similar time period and their results therefore provide our preferred benchmark for the estimation results.

Equations (2.3) to (2.9) feature a set of exogenous disturbances to technology, preferences and prices, $\{a_t, \tilde{a}_t, g_t, \tilde{p}_t, s_t, \nu_t, \mu_t\}$, which we specify to follow autoregressive processes in logs that are subject to stochastic shocks $\{\epsilon_t^a, \epsilon_t^{\tilde{a}}, \epsilon_t^g, \epsilon_t^{\tilde{p}}, \epsilon_t^s, \epsilon_t^\nu, \epsilon_t^\mu\}$. The shocks are normally distributed with mean zero and standard deviations denoted by $\{\sigma_a, \sigma_{\tilde{a}}, \sigma_g, \sigma_{\tilde{p}}, \sigma_s, \sigma_\nu, \sigma_\mu\}$. The processes for g_t, s_t and $\tilde{p}_t$ have deterministic means different from 1 that are parametrized as g, s, and $\tilde{p}$, and which will be calibrated to match business cycle moments of the steady state model. We specify autoregressive processes of order one for all shock processes, but allow the log of the commodity price $\tilde{p}_t$ to follow an AR(2). This enables us to calibrate the parameters

[19]In Appendix 2.7.5 we formally illustrate this idea in a simple setting that gives rise to the postulated relation.

[20]Akinci (2017), for example, generates a countercyclical country risk premium by introducing financial frictions in the spirit of Bernanke et al. (1999) to the economy's firm sector. Her model does not feature a commodity sector, but an extension to include it seems natural. Fernández et al. (2015) allow *future* commodity prices to affect the spread. In justifying their modeling assumptions regarding the relation between spreads and commodity prices, they make very similar arguments to the ones we have provided here.

to the ones obtained from the SVAR analysis in Section 2.2.3. The processes are

$$\begin{aligned}
ln(a_t) &= \rho_a ln(a_{t-1}) + \epsilon_t^a && (2.10)\\
ln(\tilde{a}_t) &= \rho_{\tilde{a}} ln(\tilde{a}_{t-1}) + \epsilon_t^{\tilde{a}} && (2.11)\\
ln\left(\frac{g_t}{g}\right) &= \rho_g ln\left(\frac{g_{t-1}}{g}\right) + \epsilon_t^g && (2.12)\\
ln\left(\frac{s_t}{s}\right) &= \rho_s ln\left(\frac{s_{t-1}}{s}\right) + \epsilon_t^s && (2.13)\\
ln(\nu_t) &= \rho_\nu ln(\nu_{t-1}) + \epsilon_t^\nu && (2.14)\\
ln(\mu_t) &= \rho_\mu ln(\mu_{t-1}) + \epsilon_t^\mu && (2.15)
\end{aligned}$$

and

$$ln\left(\frac{\tilde{p}_t}{\tilde{p}}\right) = \rho_{\tilde{p}}^1 log\left(\frac{\tilde{p}_{t-1}}{\tilde{p}}\right) + \rho_{\tilde{p}}^2 log\left(\frac{\tilde{p}_{t-2}}{\tilde{p}}\right) + \epsilon_t^{\tilde{p}}. \qquad (2.16)$$

The model features the following resource constraints. In the final good sector the resource constraint is given by

$$Y_t = C_t + I_t + S_t + \frac{\phi}{2}\left(\frac{K_{t+1}}{K_t} - g\right)^2 + TB_t \qquad (2.17)$$

where TB_t denotes the trade balance in final goods. The commodity market resource constraint reads as

$$\tilde{p}_t\tilde{Y}_t = \tilde{p}_t\tilde{M}_t + \tilde{TB}_t, \qquad (2.18)$$

where $\tilde{TB}_t$ denotes the real commodity trade balance, that is, net exports of commodities measured in terms of final goods. Carrying out some further national accounting, we compute the GDP and the total trade balance of the economy, both measured in terms of final goods, as

$$\begin{aligned}
Y_t^{GDP} &= Y_t + \tilde{p}_t\tilde{Y}_t - \tilde{p}_t\tilde{M}_t && (2.19)\\
TB_t^{Total} &= TB_t + \tilde{TB}_t. && (2.20)
\end{aligned}$$

The complete list of optimality conditions derived in this model is provided in Appendix 2.7.3. The Appendix also contains the derivation of a normalized version of the model that is stationary, that is, where all variables that grow in equilibrium are divided by X_{t-1}. This results in a stationary system in normalized variables, which we denote with lower case letters, and which we solve numerically with standard perturbation techniques. We carry out both a calibration exercise and a structural estimation of the model in order to asses the quantitative contribution of different shocks to fluctuations in the main macroeconomic aggregates.

2.4 Calibration and Business Cycle Characteristics

The goal of this section is to study the business cycle characteristics of the model that are induced by shocks to the commodity price. To do so, we calibrate all structural parameters of the model, including the parameters governing the stochastic process of $ln(\tilde{p}_t)$. We then generate impulse response functions, focusing exclusively on commodity price shocks.[21]

2.4.1 Calibration

Table 2.3 summarizes our baseline calibration. Many of the parameter values are standard in business cycle research, but several are worth highlighting. Both the mean of the commodity sector productivity $\tilde{a}_t$ as well as the steady state relative price of commodities $\tilde{p}$ can be adjusted to determine the relative size of the two sectors in the economy. We have normalized the mean technology in both sectors to 1 - as can be seen in equations (2.10) and (2.11) - and find the value of $\tilde{p}$ that matches the ratio of net exports of commodities to GDP observed in Argentine data (8.60%).[22] This pins down the relative size of the commodity price sector that is in line with Argentine data. The parameter d^* in equation (2.9) is calibrated to match the average trade balance to output ratio in the data (-0.041%, consistent with Table 2.1). We calibrate the mean of the exogenous spending process s to match the average government spending to GDP ratio observed in the data (9.38%). The parameter ξ, which governs the sensitivity of the interest rate spread to commodity prices, is calibrated to the value obtained from the regressions in Section 2.2.4. To be conservative, we take the lower bound of -0.199 among the statistically significant estimates we have obtained across a broad range of regression specifications. The average technology growth rate of the economy g is set directly to 1.0117 in order to generate the observed mean output growth in the data. We impose equal capital shares in both sectors ($\alpha_k = \tilde{\alpha}_k$) and set the commodity share in the final goods production to $\alpha_m = 0.05$ following Shousha (2016). The parameter ψ is typically positive but close to zero in the small open economy literature (see e.g. Schmitt-Grohe and Uribe, 2003). The estimation results of García-Cicco et al. (2010), however, highlight that the data support a larger value of this parameter. In particular, a large value is necessary to generate a standard deviation of the trade balance roughly as big as the one of output growth and a decreasing autocorrelation function of the trade balance. We therefore set $\psi = 2.8$ in line with their posterior estimate.[23] We set the adjustment cost parameter to $\phi = 6$, slightly higher than in one-sector models in the literature because this reduces the impact response of the economy to commodity shocks, which is needed to match our SVAR results (lower values would overstate

[21] We provide impulse responses functions to all other shocks in Appendix 2.7.4.

[22] To compute this target ratio in the data, we use a broad measure of commodity exports which includes manufactures of commodities. Due to data availability we use an annual sample starting in 1980.

[23] In our estimation exercise we proceed similar to García-Cicco et al. (2010) and estimate ψ.

the effect of commodity prices).[24] The stochastic process of $ln(\tilde{p}_t)$ is calibrated to be in line with the estimated SVAR coefficients in Section 2.2.3, which gives $\rho^1_{\tilde{p}} = 0.95$, $\rho^2_{\tilde{p}} = -0.13$, and $\sigma_{\tilde{p}} = 0.1064$.

Table 2.3: Model Calibration

Parameter	Value	Calibration target/source
$\tilde{p}$	0.5244	Target commodity net exports to GDP in the data (8.60%)
d^*	-0.001	Target trade balance to GDP in the data (-0.041%)
s	0.0189	Target gov't spending to GDP in the data (9.38%)
ξ	-0.199	Estimated coefficient in Section 2.2.4
g	1.0117	Average GDP growth in the data
ψ	2.8	Estimate of García-Cicco et al. (2010)
α_k	0.32	García-Cicco et al. (2010)
α_m	0.05	Shousha (2016)
$\tilde{\alpha_k}$	0.32	Impose equal capital share across both sectors
δ	0.1255	García-Cicco et al. (2010)
ϕ	6	Roughly match impact responses in SVAR
β	0.93	Steady state interest rate $\approx$ 10%
γ	2	Standard value in business cycle analysis
θ	1.6	$N^1 + N^2 \approx 1/3$
ω,$\tilde{\omega}$	1.6	Standard in SOE literature
$\rho^1_{\tilde{p}}$	0.95	Estimated SVAR coefficient (Section 2.2.3)
$\rho^2_{\tilde{p}}$	-0.13	Estimated SVAR coefficient (Section 2.2.3)
$\sigma_{\tilde{p}}$	0.1064	Estimated SVAR coefficient (Section 2.2.3)

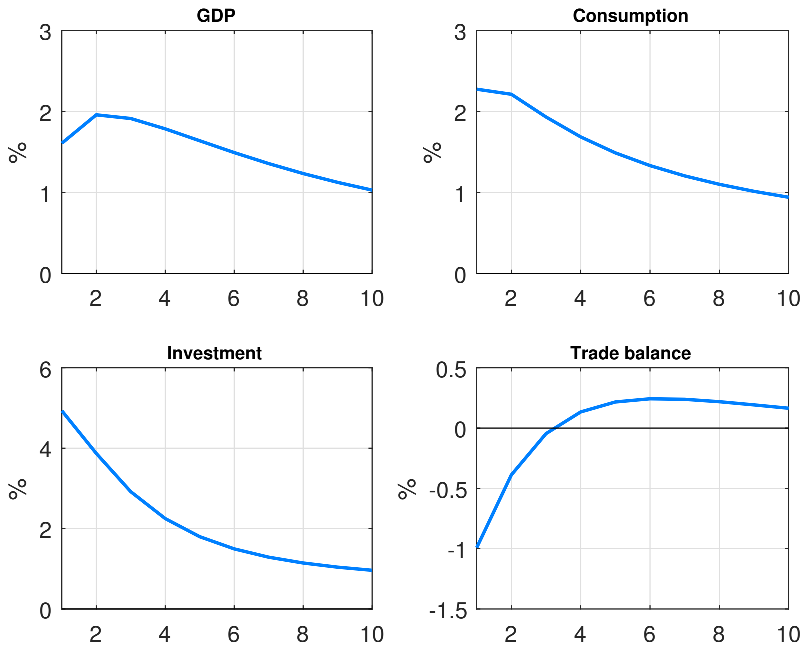
2.4.2 Impulse Response Functions to Commodity Price Shocks

Figure 2.4 displays the impulse response functions to a one-standard deviation commodity price shock $\epsilon^{\tilde{p}}_t$, using the calibration described above. The figure shows that the responses on impact are in line with the stylized facts of the business cycle of Argentina highlighted in Section 2.2. Positive commodity price shocks boost the economy by increasing total output, consumption, and investment. The investment response is the strongest, and the consumption response is larger in magnitude than the output response. The total trade balance response is negative, rendering total net exports countercyclical.

To understand the mechanism behind the dynamics visible in Figure 2.4, note that commodity prices in the model give rise to two effects. The first effect goes through commodity trade revenues. The economy needs to trade off the cost of more expensive commodity inputs in the production of final goods with the benefits of being able to produce and export commodities at higher prices (thus generating

[24]Note that the literature in general gives little guidance on sensible values for ϕ.

Figure 2.4: Impulse response functions to commodity price shock

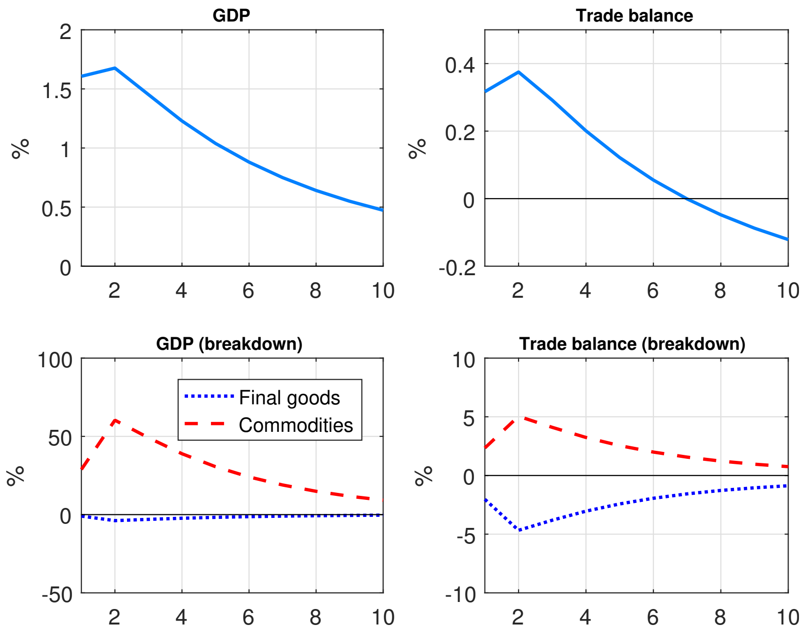


Note: Model impulse response functions to a one-standard deviation commodity price shock $\epsilon_t^{\tilde{p}}$, using the calibration described in the text.

trade revenues). The second effect is governed by the negative sensitivity of the interest spread $r_t - r^*$ to commodity prices present in equation (2.9) and based on the empirical evidence in Section 2.2.4. Both of these effects are necessary to generate the responses in Figure 2.4. To highlight this, Figures 2.5 and 2.6 open up the double role of commodity prices in our model, by plotting impulse response functions for the two effects separately and inspecting them across the two sectors of the economy. In both cases, the responses of consumption and investment growth are omitted.

Figure 2.5 studies the first effect of commodity price shocks, which we dub "competitiveness effect." The figure plots the responses of GDP and the total trade balance to a commodity price shock when setting $\xi = 0$, that is, shutting off the channel through the interest rate, which we will analyze separately below. It also breaks down these responses into the dynamics in both sectors, that is, the final good sector and the commodity sector, separately. What the left panels of the figure reveal is that after a commodity prices increase, the value-added in the commodity sector increases significantly, as higher international prices make it attractive to increase production and exports. The final good sector actually suffers, as intermediate commodity inputs necessary to produce final goods become more expensive. This effect, however, is dwarfed by the boom in the commodity sector and total production in the economy increases. The trade balances in the two sectors, shown in the right panels of the figure, move in different directions. The economy starts exporting more commodities and importing final goods, as the former are very attractive to sell abroad and the latter less attractive to produce domestically. Looking at the two sectors together, the *total* trade surplus increases with the commodity price increase. This highlights that the first effect alone does not generate a countercyclical *total* trade balance, which is a salient feature in emerging economy business cycle data.

Figure 2.5: Breakdown of IRFs: No interest rate channel ($\xi = 0$)



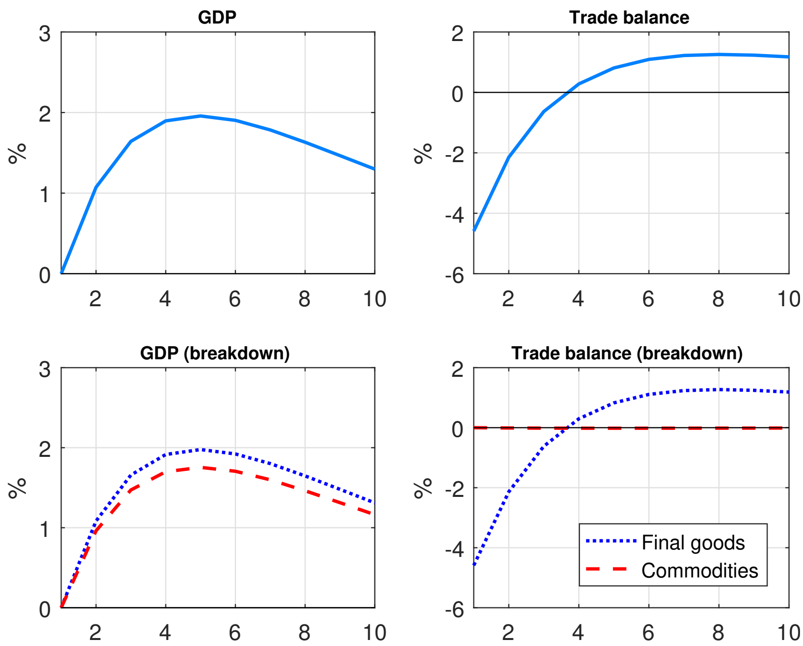
Note: Model impulse response functions to a one-standard deviation commodity price shock $\epsilon_t^{\tilde{p}}$, using the calibration described in the text but setting $\xi = 0$.

Figure 2.6 shows the dynamics arising from the second effect, which we call "borrowing cost effect." The figure plots the IRFs of total GDP and the total trade balance to a simple interest rate shock. This shock is (qualitatively) isomorphic to an increase in commodity prices that only goes through the presence of $\tilde{p}_t$ in equation (2.9) but that does not *directly* affect production in either sector.[25] It thus completely shuts off the competitiveness channel described above and only shows the effect that commodity price have through the spread between the economy's borrowing rate and the world interest rate. As before, the figure breaks down the response by displaying the dynamics in each sector separately. The figure shows that the exogenous fall in borrowing rates allows households and firms to bring resources to the present by borrowing funds and decreasing the final good trade balance, that is, importing final goods. Some of these resources will be consumed (consumption goes up on impact, not shown in the figure), and some will be invested into capital (investment goes up on impact, not shown in the figure) in order to produce final goods and maintain a smooth path of consumption. Some of the capital will also be used to produce commodities, which are a required intermediary input to final good production. This gives a slow and hump-shaped increase in the GDP of each sector

[25]For the purpose of the comparison, the standard deviation of the interest rate shock is calibrated to have the same maximum output response as the total response in Figure 2.4. The persistence is set to 0.9.

and of the total economy. Hence, the total trade balance falls and output increases, but not on impact. This lack of impact response in output stands in contrast with the empirical impulse responses and suggests that this channel alone cannot mimic the data.

Figure 2.6: Breakdown of IRFs: Pure interest rate shock



Note: Model impulse response functions to a one-standard deviation interest rate shock ϵ_t^{μ}, using the calibration described in the text.

In conclusion, the double-role of commodity prices in our model, through the *joint* impact of the competitiveness and the borrowing cost channels, gives rise to dynamics that are well in line with the empirical regularities observed in Argentina, as shown by comparing the SVAR results from Figure 2.3 with the model responses presented in Figure 2.4. This insight further highlights the importance of endogenously countercyclical spreads for aggregate fluctuations in commodity exporting economies, as recently also noted by Fernández et al. (2015) and Shousha (2016).[26]

We emphasize again that the focus of the calibration exercise in this section lies on explaining the dynamics that arises from commodity shocks alone. This is done to highlight our mechanism in light of the facts present in the data.[27] In order to systematically gauge the fraction of aggregate fluctuations that can be accounted for by commodity price shocks, in comparison to all other shocks, we move on to estimating the model in the next section.

[26]Recent work by Ben Zeev et al. (2017) and Farias and da Silva (2017) focuses on commodity price news shocks. These news shocks might also be connected with (and in fact be capturing) the interest rate effect, a link that deserves further exploration.

[27]In Appendix 2.7.4 we report the IRFs to all of the other shocks we have defined in the model.

2.5 Estimation: Assessing the Quantitative Contribution of Different Sources of Shocks in Emerging Economies

In this section our goal is to assess the quantitative contribution of different shocks to aggregate fluctuations in emerging economies for which commodity exports are potentially important. To do so, we take the model to Argentine data and structurally estimate it with the goal of running a "horse race" between the various shocks that possibly drive the business cycle. We maintain the calibration of most of the parameters (see Table 2.3), and estimate the stochastic processes of the exogenous disturbances defined by equations (2.10) to (2.16). In addition, we also estimate two key structural parameters. The first is at the heart of our mechanism: ξ, which governs the sensitivity of the real interest rate spread to commodity prices. Estimating this parameter allows the data to speak about the strength of this mechanism within our model structure. Furthermore, we estimate ψ, a parameter that governs the trade balance dynamics in the economy.[28] In carrying out the estimation exercise, we give equal footing to all different shocks, which correspond to the candidate triggers previously proposed in the literature.

2.5.1 Estimation Specification

We carry out a Bayesian estimation defining standard priors on the estimated parameters. We run a Markov Chain Monte Carlo (MCMC) algorithm to obtain draws from the marginal posterior distributions of the parameters.[29] We then compute forecast error variance decompositions as well as historical variance decompositions of the observables at the estimated posterior modes. To estimate the model we add the following measurement equations

$$\begin{aligned} \Delta lnY_t^{GDP,obs} &= lnY_t^{GDP} - lnY_{t-1}^{GDP} && (2.21)\\ \Delta lnC_t^{obs} &= lnC_t - lnC_{t-1} && (2.22)\\ \Delta lnI_t^{obs} &= lnI_t - lnI_{t-1} && (2.23)\\ TB^{Total,obs}/Y_t^{GDP,obs} &= TB_t^{Total}/Y_t^{GDP}, && (2.24) \end{aligned}$$

where $\Delta lnY_t^{GDP,obs}, \Delta lnC_t^{obs}, \Delta lnI_t^{obs}$ and $\Delta TB^{Total,obs}$ correspond to the empirically observed time series which we analyzed in Section 2.2.[30] The variables on the right hand side of equations (2.21) to (2.24) are model concepts defined in Section 2.3.[31] As explained above, we estimate the parameters governing the stochastic

[28] The importance of estimating this parameter has been stressed by García-Cicco et al. (2010).

[29] We take 10 million draws. We discard the first 25% of draws and keep the remaining ones for inference. The acceptance ratio is 27.3%.

[30] In principle we could add the commodity price series, which we used for parts of the calibration of the model, as an observable. However, since the Grilli and Yang (1988) may capture some dynamics unrelated to prices actually faced by Argentina, and an Argentina-specific index is only available for a much shorter sample, our preferred specification is to estimate the model without this observable and then compare the model-implied commodity price process with the empirically observed index. See the discussion further below.

[31] Note that while we solve the (linearized) model in variables that are normalized by X_{t-1} (see

processes of all shocks, as well as ξ and ψ (all other parameters are calibrated as before). Table 2.4 summarizes the priors imposed on the parameters. As is standard in the estimation of DSGE models, we use beta priors on the persistence parameters and inverse-gamma priors on the standard deviations. The parameter values of the priors are the same as in Smets and Wouters (2007) and a number of related papers, except for the commodity price process. Since the latter is specified as an AR(2), we use priors that at the mode impose the same maximum root as for the other disturbances.[32] We set identical scale parameters on the standard deviation of the shocks to remain agnostic about the relative importance of the different shocks. We put a normal prior on ξ, which is centered around the smallest statistically significant regression estimate from Section 2.2.4, with the standard deviation equal to the standard error obtained from the regression. Finally, our prior on ψ, also normal, is centered around the estimate obtained by García-Cicco et al. (2010).

Table 2.4: Estimated parameters and priors

Parameter	**Prior**	**Mean**	**Std. dev.**
ξ	Normal	-0.199	0.045
ψ	Normal	2.8	0.5
$\rho_{\tilde{p}}^1$	Beta	0.8	0.2
$-\rho_{\tilde{p}}^2$	Beta	0.15	0.1
$\sigma_{\tilde{p}}$	Inverse-Gamma	0.05	2
ρ_i	Beta	0.5	0.2
σ_i	Inverse-Gamma	0.05	2
$i = a, \tilde{a}, g, s, \nu, \mu$			

2.5.2 Estimation Results

How large is the contribution of different structural shocks to the variation in output, consumption, investment and the trade balance in emerging economies? We address this question using the results in Table 2.5. Panel (a) of the table shows the results of an (infinite horizon) forecast error variance decomposition based on the posterior estimates of our model using Argentine data from 1900 through to 2015.[33] For each of the variables used as observables, this gives the share of variation that can be explained by a particular shock. We begin by focusing on the commodity price shock, as this is the main difference with respect to Aguiar and Gopinath (2007) and

Appendix 2.7.3), we here use growth rates in the original non-normalized variables. This is possible, as the implied nonstationary variables can be recomputed from the model solution.

[32] $\rho_{\tilde{p}}^1 = 0.8$ and $\rho_{\tilde{p}}^2 = -0.15$ imply that the larger root of the process 0.5, which is the same for an AR(1) processes with $\rho = 0.5$.

[33] Table 2.8 in the appendix reports posterior mean and credible intervals of the individual parameters we estimate.

García-Cicco et al. (2010). As the table reveals, a sizable fraction of output (21.67%), consumption (24.02%) and investment growth (34.11%) can be explained by commodity price shocks. This confirms the intuition we derived from the calibration exercise and from the responses that were present in our SVAR analysis.

Table 2.5: Variance decomposition for baseline estimation

	Stationary technology	Nonstat. technology	Interest rate	Comm. price	Spending shock	Pref. Shock
			(a) Baseline sample from 1900-2015			
Output growth	51.15%	20.55%	1.12%	**21.67%**	0.19%	5.33%
Consumption growth	35.32%	10.87%	3.24%	**24.02%**	1.51%	25.05%
Investment growth	11.68%	2.15%	23.8%	**34.11%**	1.9%	26.35%
Trade balance	1.19%	2.53%	64.71%	**16.33%**	2.08%	64.71%
			(b) Shorter sample from 1950-2015			
Output growth	39.14%	20.57%	0.69%	**37.97%**	0.08%	1.54%
Consumption growth	28.47%	11.72%	2.01%	**42.28%**	1.14%	14.39%
Investment growth	9.48%	2.57%	15.35%	**61.11%**	0.50%	10.99%
Trade balance	1.28%	3.03%	52.83%	**31.56%**	0.42%	10.87%

Note: Forecast error variance decomposition (at infinite horizon) of the observables used for estimation, calculated at the posterior modes. Stationary technology is the sum of the contribution of a_t and $\tilde{a}_t$. These estimates are obtained from the baseline estimation specification explained in the text.

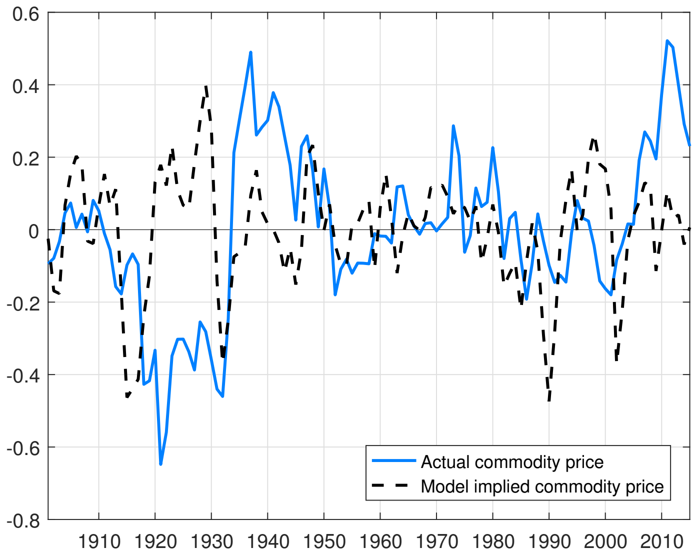
Turning to the other shocks, the table shows that our estimation attributes most of the variation in output growth (51.15%) to transitory technology shocks (the table reports the joint contribution of a_t and $\tilde{a}_t$). This finding is in line with García-Cicco et al. (2010). We do not, however, confirm their conclusion regarding the very small contribution of shocks to nonstationary technology à la Aguiar and Gopinath (2007). We find the contribution of these shocks to be sizable, explaining 20.55% of the variation in output growth in Argentina.[34] Preference shocks and interest shocks also play an important role in understanding the business cycle. The former, affecting directly the intertemporal choices of the household, explains in particular consumption and investment growth, as well as trade balance variation, while the latter also contributes substantially to the variance of investment growth. The government spending (endowment) shock is generally found to be unimportant, which is in line with the previous literature.

To shed further light on our findings with respect to commodity prices, in Figure 2.7 we plot two series. The first one, indicated with the dashed black line, corresponds to the model-implied commodity price process, that is, the time series of $\tilde{p}_t$ obtained from feeding the estimated shocks $\epsilon_t^{\tilde{p}}$ into equation (2.16) and setting

[34] Interestingly, Akinci (2017) also finds both types of technology shocks to be important in the context of a model that features financial frictions and time-varying risk premia. This is in contrast with Chang and Fernandez (2013), who find that nonstationary productivity shocks play a minor role relative to stationary TFP and interest shocks, broadly confirming the results of García-Cicco et al. (2010). None of these studies feature a role for commodity prices.

the parameters $\rho^1_{\tilde{p}}$ and $\rho^2_{\tilde{p}}$ to their estimated posterior mode. The second series, indicated with a solid blue line, shows the real commodity price index, which we have plotted and used for calibrating parts of the model above. It is apparent that, reassuringly, the two time series broadly share common features, such as a similar volatility and reasonably synchronized movements. This is particularly the case in the post-1950 period, while the war and interwar period give rise to some large level differences between the two price series. The wars are special periods in which trade barriers an production are affected, giving room to large swings in trade and commodity prices that were not connected in the way our theory would prescribe. (Trade barriers fluctuated significantly during this period, opening a volatile gap between international commodity prices and the actual prices received by Argentine producers.) Furthermore, we point out that the commodity price index by Grilli and Yang (1988) captures world commodity prices and not necessarily those commodity prices faced by Argentina. With growing financial integration, the global cross-section of commodity prices has become more correlated over time and thus may render the index more closely related to the actual commodity prices faced by Argentina in the later parts of the estimation sample.

Figure 2.7: Estimated and actual process for commodity prices



Note: The blue solid line repeats the commodity price series from Figure 2.2. The dashed black line is the commodity price process $\tilde{p}_t$ that is implied by the posterior estimates of the parameters and shocks of the estimated model.

Given these concerns, we re-estimated the model using a subsample of the data from 1950 to 2015. The results of the forecast error variance decomposition are shown in Table 2.5, Panel (b). In this sample, the quantitative contribution of commodity price shocks is estimated to be even larger. Commodity price shocks explain 37.97% of the variance in output growth, 42.28% in consumption growth and 61.11% in investment growth. The relative importance of other shocks remains broadly similar in this sample.

While we primarily focus on comparing our quantitative results to García-Cicco et al. (2010), as these authors use a similarly long sample for Argentina, our findings are also broadly in line with comparable recent work on commodity price shocks in emerging markets that has estimated quantitative models on shorter samples. Fernández et al. (2015), for example, estimate that the share of commodity shocks in the variance of real output across a number of emerging economies is 42%, a number that is very similar to our post-1950 estimate.

In addition to the decomposition given in Table 2.5, which is a theoretical object computed at the posterior modes, it is also possible to construct a *historical* variance decomposition that breaks down the movements of a variable *at a given point in the actual data sample* into the contribution of the different shocks. Figure 2.8 presents such a decomposition for Argentine output growth from 1900 to 2015. The black line displays the actual time series of growth in real GDP per capita, which is used as one of the observables in the estimation. The bars represent the contribution of different shocks to the movements in the output time series at given points in time. Overall, the figure mirrors the insights from Table 2.5, given that commodity price shocks and technology shocks (of both types), capture most of the variation in output growth. Figure 2.8, in addition, enables us to inspect particular episodes in the economic history of Argentina, as scrutinized for example by Taylor (2014), and interpret them through the lens of our model.

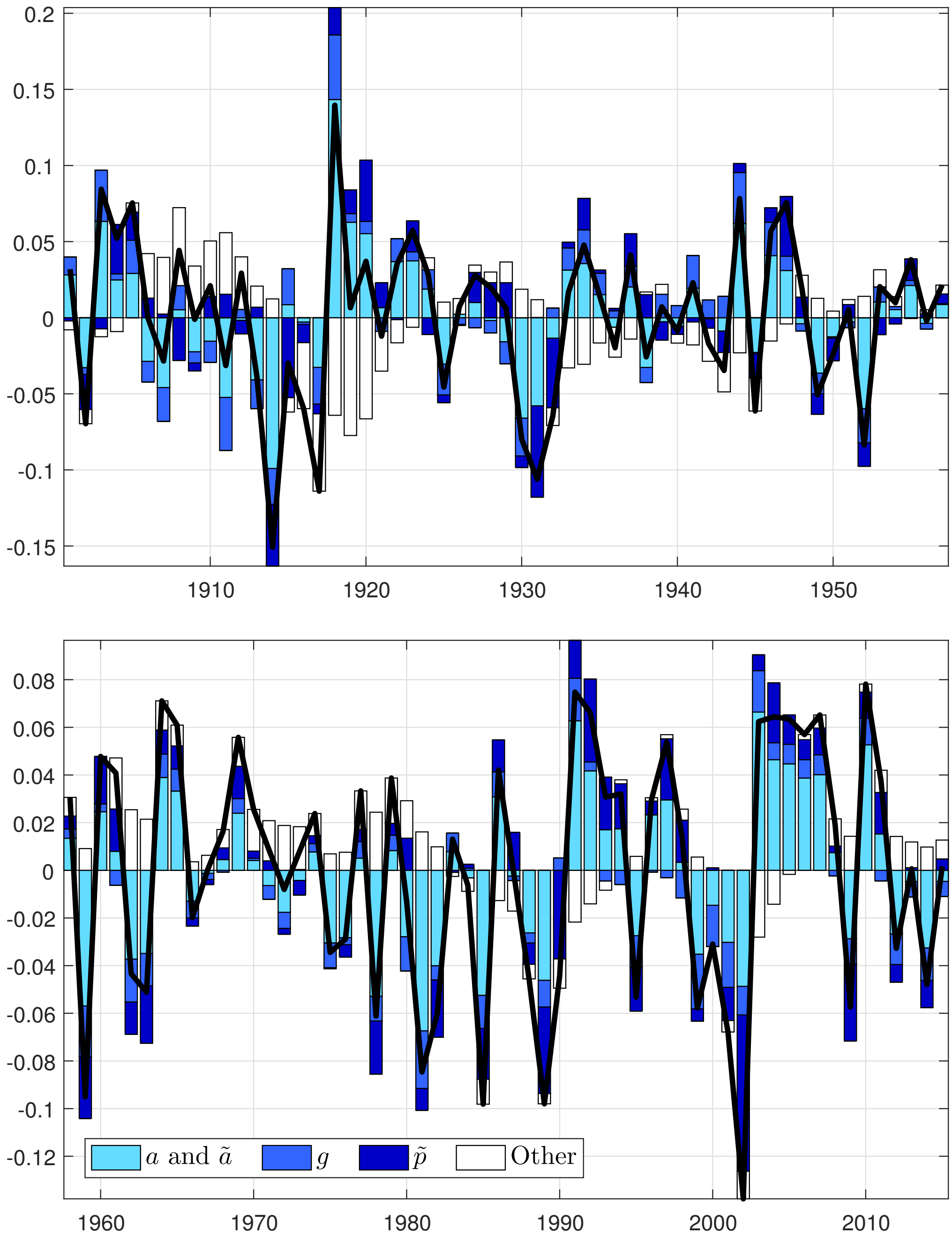
Taken altogether, our results suggest that commodity prices should feature prominently in the analysis of business cycles in emerging economies. In terms of quantitative contribution, we find that they are among the three most important shocks in driving output growth in Argentina. Importantly, shocks to international commodity prices, in contrast to inherently very different concepts such as domestic TFP shocks, are easier to measure and identify, and eventually act upon, by policy makers.

2.5.3 Further Discussion: Measurement of GDP

How "direct" is the effect of commodity price variation on real GDP? The relative price $\tilde{p}_t$ directly enters the calculation of real GDP in our model, but national accounting techniques in practice may not reflect the full variation in relative prices in the way our measurement equation (2.19) prescribes.[35] It is therefore of interest to break down the variation in GDP resulting from commodity price fluctuations into the share that comes directly from $\tilde{p}_t$ and the share that arises from the endogenous changes in quantities following commodity price changes. This latter effect on quantities would be the only source of change in measured real GDP if the statistical office kept prices constant in its measurement. If this share of the variation is important, then the effect of commodity price shocks on GDP that we measure would be more robust to the specific measurement of real GDP in practice.

[35] This could be due to base-year pricing, chain-linking or simply due to price mismeasurement or interpolation. Kehoe and Ruhl (2008), for example, argue that changes in the terms of trade have no first-order effect if output is measured as chain-weighted real GDP.

Figure 2.8: Historical Decomposition of Argentine Output Growth 1900-2015



Note: The line displays the actual time series of real GDP per capita growth, which is used as one of the observables in the estimation. The bars represent the contribution of different shocks to the movements in this series at a given point in time. The estimates are obtained using the baseline estimation specification explained in the text. "Other" includes ν, s, μ and the contribution of initial values, which is negligible.

To study this question, Figure 2.10 in the Appendix plots two alternative measures of GDP from a simulation exercise. The exercise consists of feeding observed commodity prices into the model, holding all other disturbances constant, and then computing two alternative GDP measures. The first measure is Y_t^{GDP} computed as in equation (2.19), whereas the second one, $Y_t^{GDP^*}$, computes the economy's GDP holding commodity prices fixed at their steady state value $\tilde{p}$, that is,

$$Y_t^{GDP^*} = Y_t + \tilde{p}\tilde{Y}_t - \tilde{p}\tilde{M}_t. \tag{2.25}$$

The figure shows that the two resulting series are very similar, and the variation in $Y_t^{GDP^*}$ accounts for most of the variation in Y_t^{GDP}.[36] This highlights that the economy's endogenous dynamics in response to changes in international commodity prices accounts for the major bulk in the variation of total real value added. This makes the results in our paper robust to different methods used to measure real GDP.

2.6 Conclusion

This paper has sought to answer a classical question in International Macroeconomics: what causes the large swings in economic activity in emerging markets? The literature has proposed a variety of triggers, but remains split on the answers. We study the question anew, combining a model that nests the previous sources of shocks advanced in the literature and historical data for Argentina going back to 1900.

The model features two key elements. First, it allows for a second sector to capture the separate role of commodities in the economy. Specifically, the analysis focuses on the case of a net commodity exporting country, facing exogenous price changes. Second, the model embeds a negative relation between the interest rate premium and commodity prices, which is consistent with the empirical evidence. Exogenous increases in commodity prices improve both the competitiveness of the economy and its borrowing terms through the negative effect of higher prices on the spread between the country's borrowing rates and world interest rates. Both effects jointly result in strongly positive effects of commodity price movements on GDP, consumption, and investment, and a negative effect on the total trade balance. They also generate an excess response of consumption over output.

We estimate the model using data on Argentina from 1900 to 2015 to provide a quantitative evaluation of the various sources of shocks and their effect on macroeconomic aggregates. Our estimate of the contribution of commodity price shocks to fluctuations in output growth of Argentina is in the order of 22%. Furthermore, commodity prices account for 24% and 34% of the variation in consumption and investment growth, respectively. The contribution of these shocks is even bigger on a post-1950 data sample, accounting for 38% of the variance of output growth,

[36] The R-squared from regressing one series on the other is 0.95.

42% of consumption, and 61% of investment. We also find a role for non-stationary productivity shocks - albeit much smaller than the one documented in Aguiar and Gopinath (2007), though bigger than García-Cicco et al. (2010) - and an important role for stationary productivity shocks, consistent with previous findings.

Though in this paper we do not address normative issues, the results offer hope. Insofar as part of the cycle can be accounted for by observable variables (international commodity prices) that cannot be manipulated for political goals, contingent macroeconomic policies can be designed to help mitigate the cycle. Given the nature of the driver, sovereign wealth funds may offer a promising avenue for tackling volatility in commodity producing countries like Argentina. A proper normative analysis would require, at a minimum, an extension of the model to incorporate default, a task we leave for future work.

2.7 Appendices

2.7.1 Details on Data

GDP and its components

Data on real GDP, Investment, Consumption, Government Spending and Net Exports from 1900 through to 2009 come from Ferreres (2005) - Ferreres has extended these series to 2009. We extend the data further to 2015 using the corresponding series from the Argentine Finance Ministry "Ministerio de Economia (Ejecución Presupuestaria de la Administración Nacional)," available online. The growth rate of the latter series was applied to Ferreres' 2009 figure.

Commodity Prices

Data on world commodity prices are based on the Grilli and Yang (1988) commodity price index series updated by Pfaffenzeller et al. (2007), which runs from 1900 through to 2011. We update the series to 2015, following Pfaffenzeller et al. (2007)'s procedure.

The Argentina-specific price index is constructed using Argentine export weights available in the UN Comtrade data base. We match these weights with commodity-specific price indeces provided by the World Bank. This is done for the broad commodity categories fuel, timber, food, beverages and fertilizer from 1962.

As a deflator for the commodity price series we use the index of US-dollar import prices for Argentina provided by Pfaffenzeller et al. (2007), which we update till 2015 using the figures from INDEC. For robustness we also tried manufacturing prices (expressed in US dollars), and the US consumer price index, available via FRED. The results remain broadly unchanged using these deflators.

World Real Interest Rate

To measure global real interest rates we use the UK nominal interest rate series published by the Bank of England from 1900 through 2015 and subtract the UK inflation rate provided by the UK Office for National Statistics (ONS).

Domestic Real Interest Rates

We use the nominal domestic lending rate, savings rate and money market rate, provided by the IMF International Financial Statistics. We deflate these series using the corrected inflation measure available at http://www.inflacionverdadera.com/. See Cavallo (2013) for a discussion.

Government Debt

Data on Debt-to-GDP ratios come from Argentina's national statistical office, INDEC (Online, Table 7.10).

2.7.2 Additional Regression Results

Table 2.6: Additional Regression Results: Using the Lending Rate

	(1)	(2)	(3)	(4)	(5)
LHS variable	Real spread (**based on savings rate**)				
Commodity price	-0.131	-0.123	-0.174	-0.138	-0.188
	(0.111)	(0.113)	(0.117)	(0.116)	(0.119)
Output growth		-0.317			-0.259
		(0.426)			(0.427)
Trade balance			-0.526		-1.398
			(0.478)		(0.906)
Debt-to-GDP ratio				-0.020	0.154
				(0.075)	(0.139)
Constant	-0.113***	-0.107***	-0.106***	-0.102*	-0.176**
	(0.026)	(0.027)	(0.026)	(0.050)	(0.075)
Observations	25	25	25	25	25
R-squared	0.057	0.080	0.106	0.060	0.183

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The real spread is calculated by deflating the domestic savings rate, provided by the IMF, with a corrected inflation measure (see Cavallo, 2013), and then subtracting the UK real rate. The commodity price is in log deviations from mean, as plotted in Figure 2.2, Panel (b). Appendix 2.7.1 provides details on the sources of the other regressors.

Table 2.7: Additional Regression Results: Using the Money Market Rate

	(1)	(2)	(3)	(4)	(5)
LHS variable	Real spread (**based on money market rate**)				
Commodity price	-0.183	-0.165	-0.175	-0.162	-0.178
	(0.187)	(0.184)	(0.206)	(0.196)	(0.207)
Output growth		-0.941			-0.931
		(0.641)			(0.661)
Trade balance			0.088		-0.579
			(0.829)		(1.377)
Debt-to-GDP ratio				0.052	0.107
				(0.122)	(0.203)
Constant	0.031	0.044	0.030	0.003	-0.004
	(0.038)	(0.039)	(0.042)	(0.078)	(0.102)
Observations	34	34	34	34	34
R-squared	0.029	0.092	0.029	0.035	0.101

Standard errors in parentheses
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Note: The real spread is calculated by deflating the money market rate, provided by the IMF, with a corrected inflation measure (see Cavallo, 2013), and then subtracting the UK real rate. The commodity price is in log deviations from mean, as plotted in Figure 2.2, Panel (b). Appendix 2.7.1 provides details on the sources of the other regressors.

2.7.3 Model Details

Optimality conditions

Firms

The first-order conditions for final goods producers with respect to K_t^1, N_t^1 and $\tilde{M}_t$ are

$$\begin{aligned} r_t^{k1} &= \alpha_K a_t (K_t^1)^{\alpha_K - 1} (\tilde{M}_t)^{\alpha_M} (X_t N_t^1)^{1-\alpha_K-\alpha_M} && (2.26) \\ w_t^1 &= (1-\alpha_K-\alpha_M) a_t (K_t^1)^{\alpha_K} (\tilde{M}_t)^{\alpha_M} (X_t N_t^1)^{-\alpha_K-\alpha_M} X_t && (2.27) \\ \tilde{p}_t &= \alpha_M a_t (K_t^1)^{\alpha_K} (\tilde{M}_t)^{\alpha_M - 1} (X_t N_t^1)^{1-\alpha_K-\alpha_M}. && (2.28) \end{aligned}$$

The first-order conditions for commodity producers with respect to K_t^1 and N_t^1 are

$$\begin{aligned} r_t^{k2} &= \tilde{\alpha_K} \tilde{p}_t \tilde{a}_t (K_t^2)^{\tilde{\alpha_K}-1} (X_t N_t^2)^{1-\tilde{\alpha_K}} && (2.29) \\ w_t^2 &= (1-\tilde{\alpha_K}) \tilde{p}_t \tilde{a}_t (K_t^2)^{\tilde{\alpha_K}} (X_t N_t^2)^{-\tilde{\alpha_K}} X_t && (2.30) \end{aligned}$$

Representative Household

Setting up the dynamic Lagrangian

$$\begin{aligned} \mathcal{L} = \sum_{t=0}^{\infty} \nu_t \beta^t \Bigg\{ & \frac{[C_t - \theta\omega^{-1} X_{t-1}(N_t^1)^{\omega} - \theta\tilde{\omega}^{-1} X_{t-1}(N_t^2)^{\tilde{\omega}}]^{1-\gamma} - 1}{1-\gamma} \\ & - X_{t-1}^{-\gamma} \lambda_t \Bigg[C_t + K_{t+1}^1 + K_{t+1}^2 + D_t + S_t + \frac{\phi}{2}\left(\frac{K_{t+1}}{K_t} - g\right)^2 \\ & - r_t^{k1}(K_t^1) - r_t^{k2}(K_t^2) - w_t^1 N_t^1 - w_t^2 N_t^2 - (1-\delta)K_t^1 - (1-\delta)K_t^2 - \frac{D_{t+1}}{1+r_t} \Bigg] \Bigg\}, \end{aligned} \tag{2.31}$$

the first-order conditions with respect to C_t, N_t^1, N_t^2, D_{t+1}, K_{t+1}^1, and K_{t+1}^2 are derived as follows:

$$[C_t - \theta\omega^{-1} X_{t-1}(N_t^1)^{\omega} - \theta\tilde{\omega}^{-1} X_{t-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma} = \lambda_t X_{t-1}^{-\gamma} \tag{2.32}$$

$$[C_t - \theta\omega^{-1} X_{t-1}(N_t^1)^{\omega} - \theta\tilde{\omega}^{-1} X_{t-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma} \theta X_{t-1} (N_t^1)^{\omega-1} = \lambda_t X_{t-1}^{-\gamma} w_t^1 \tag{2.33}$$

$$[C_t - \theta\omega^{-1} X_{t-1}(N_t^1)^{\omega} - \theta\tilde{\omega}^{-1} X_{t-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma} \theta X_{t-1} (N_t^2)^{\tilde{\omega}-1} = \lambda_t X_{t-1}^{-\gamma} w_t^2 \tag{2.34}$$

$$\nu_t \lambda_t X_{t-1}^{-\gamma} = \beta(1+r_t) X_t^{-\gamma} \mathbb{E}_t(\nu_{t+1}\lambda_{t+1}) \tag{2.35}$$

$$\begin{aligned} & \nu_t \lambda_t X_{t-1}^{-\gamma} \left[1 + \phi\left(\frac{K_{t+1}}{K_t} - g\right) \right] = \\ & \beta X_t^{-\gamma} \mathbb{E}_t \left\{ \nu_{t+1}\lambda_{t+1} \left[r_{t+1}^{k1} + 1 - \delta + \phi\left(\frac{K_{t+2}}{K_{t+1}} - g\right)\frac{K_{t+2}}{K_{t+1}} - \frac{\phi}{2}\left(\frac{K_{t+2}}{K_{t+1}} - g\right)^2 \right] \right\} \end{aligned} \tag{2.36}$$

$$\nu_t\lambda_t X_{t-1}^{-\gamma}\left[1+\phi\left(\frac{K_{t+1}}{K_t}-g\right)\right]=$$
$$\beta X_t^{-\gamma}\mathbb{E}_t\left\{\nu_{t+1}\lambda_{t+1}\left[r_{t+1}^{k2}+1-\delta+\phi\left(\frac{K_{t+2}}{K_{t+1}}-g\right)\frac{K_{t+2}}{K_{t+1}}-\frac{\phi}{2}\left(\frac{K_{t+2}}{K_{t+1}}-g\right)^2\right]\right\} \tag{2.37}$$

Note that equations (2.36) and (2.37) imply that the expected return on capital is equalized across the two sectors in the economy.

Stationary version of equilibrium

Imposing market clearing and denoting $c_t = \frac{C_t}{X_{t-1}}$, $k_t^1 = \frac{K_t^1}{X_{t-1}}$, $k_t^2 = \frac{K_t^2}{X_{t-1}}$ etc., and using the fact that $g_t = X_t/X_{t-1}$, the first-order conditions (2.32) to (2.37) can be rewritten in stationary form as:

$$[c_t-\theta\omega^{-1}(N_t^1)^{\omega}-\theta\tilde{\omega}^{-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma}=\lambda_t \tag{2.38}$$

$$\begin{aligned}&[c_t-\theta\omega^{-1}(N_t^1)^{\omega}-\theta\tilde{\omega}^{-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma}\theta(N_t^1)^{\omega-1}\\ &=\lambda_t g_t^{(1-\alpha_K-\alpha_M)}(1-\alpha_K-\alpha_M)a_t(k_t^1)^{\alpha_K}(\tilde{m}_t)^{\alpha_M}(N_t^1)^{-\alpha_K-\alpha_M}\end{aligned} \tag{2.39}$$

$$\begin{aligned}&[C_t-\theta\omega^{-1}(N_t^1)^{\omega}-\theta\tilde{\omega}^{-1}(N_t^2)^{\tilde{\omega}}]^{-\gamma}\theta(N_t^2)^{\tilde{\omega}-1}\\ &=\lambda_t g_t^{(1-\tilde{\alpha_K})}(1-\tilde{\alpha_K})\tilde{p}_t\tilde{a}_t(k_t^2)^{\tilde{\alpha_K}}(N_t^2)^{-\tilde{\alpha_K}}\end{aligned} \tag{2.40}$$

$$\lambda_t=\beta(1+r_t)g_t^{-\gamma}\mathbb{E}_t\left(\frac{\nu_{t+1}}{\nu_t}\lambda_{t+1}\right) \tag{2.41}$$

$$\tilde{p}_t=\alpha_M g_t^{(1-\alpha_K-\alpha_M)}a_t(k_t^1)^{\alpha_K}(\tilde{m}_t)^{\alpha_M-1}(N_t^1)^{1-\alpha_K-\alpha_M} \tag{2.42}$$

$$\begin{aligned}&\nu_t\lambda_t\left[1+\phi\left(\frac{k_{t+1}}{k_t}g_t-g\right)\right]=\\ &\beta g_t^{-\gamma}\mathbb{E}_t\left\{\nu_{t+1}\lambda_{t+1}\left[g_t^{1-\alpha_K-\alpha_M}\alpha_K a_{t+1}(k_{t+1}^1)^{\alpha_K-1}(\tilde{m}_{t+1})^{\alpha_M}(N_{t+1}^1)^{1-\alpha_K-\alpha_M}\right.\right.\\ &\left.\left.+1-\delta+\phi\left(\frac{k_{t+2}}{k_{t+1}}g_{t+1}-g\right)\frac{k_{t+2}}{k_{t+1}}-\frac{\phi}{2}\left(\frac{k_{t+2}}{k_{t+1}}g_{t+1}-g\right)^2\right]\right\}\end{aligned} \tag{2.43}$$

$$\nu_t\lambda_t\left[1+\phi\left(\frac{k_{t+1}}{k_t}g_t-g\right)\right]=$$

$$\beta g_t^{-\gamma}\mathbb{E}_t\Bigg\{\nu_{t+1}\lambda_{t+1}\Bigg[g_t^{1-\tilde{\alpha_K}}\tilde{\alpha_K}\tilde{p}_{t+1}\tilde{a}_{t+1}(k^2_{t+1})^{\tilde{\alpha_K}-1}(N^2_{t+1})^{1-\tilde{\alpha_K}} \tag{2.44}$$

$$+1-\delta+\phi\left(\frac{k_{t+2}}{k_{t+1}}g_{t+1}-g\right)\frac{k_{t+2}}{k_{t+1}}-\frac{\phi}{2}\left(\frac{k_{t+2}}{k_{t+1}}g_{t+1}-g\right)^2\Bigg]\Bigg\}$$

The remaining equations of the system that define the stationary equilibrium are given by the budget constraint (with factor prices eliminated), the production functions and the interest rate equation, all normalized in the same way, i.e. by

$$c_t+k_{t+1}g_t+\tilde{p}_t\tilde{m}_t+d_t+s_t+\frac{\phi}{2}\left(\frac{k_{t+1}}{k_t}g_t-g\right)^2=y_t+\tilde{p}_t\tilde{y}_t+(1-\delta)k_t+\frac{d_{t+1}}{1+r_t}g_t \tag{2.45}$$

$$y_t=a_t(k^1_t)^{\alpha_K}(\tilde{m}_t)^{\alpha_M}(N^1_t)^{1-\alpha_K-\alpha_M} \tag{2.46}$$

$$\tilde{y}_t=\tilde{a}_t(k^2_t)^{\tilde{\alpha_K}}(N^2_t)^{1-\tilde{\alpha_K}} \tag{2.47}$$

$$r_t=r^*+\psi\big(e^{d_{t+1}-d^*}-1\big)-\xi\left(log(\tilde{p}_t)-log(\tilde{p})\right)+\big(e^{\mu_t-1}-1\big) \tag{2.48}$$

and by the stochastic processes (2.10) to (2.16) in the body of the paper. The total trade balance and GDP of the economy can be calculated accordingly.

Steady state

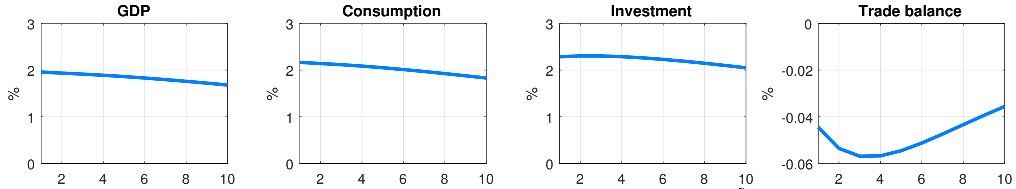
To compute the steady state, we can proceed as follows:

1. Drop all time subscripts.
2. Steady state must fulfill $r=r^*=\frac{1}{\beta}g^{-\gamma}-1$ and $d=d^*$ from (2.41) and (2.48).
3. Solve (2.44) for the steady state capital-labor ratio in the commodity sector as a function of primitives
4. Combine (2.38) and (2.39) through λ. Plug in the capital-labor ratio. It is possible to solve analytically for N^2 as a function of primitives. Using the capital-labor ratio, can solve for k^2.
5. Combine (2.38), (2.40), (2.42), and (2.43) to eliminate λ, k^1, $\tilde{m}$. Obtain an equation for N^1 as an implicit function of primitives. Solve this equation for N^1 *numerically.*
6. Use the equations combined in the previous step to solve for k^1 and $\tilde{m}$ given the solution for N^1.
7. Use the budget constraint to solve for c.

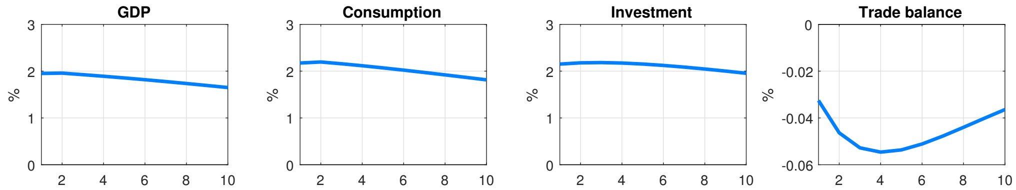
2.7.4 Additional Model Results

Figure 2.9: Impulse response functions to different shocks

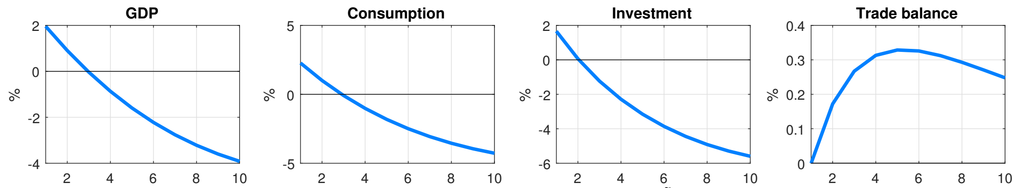
(a) Final goods sector productivity shock ϵ_t^a



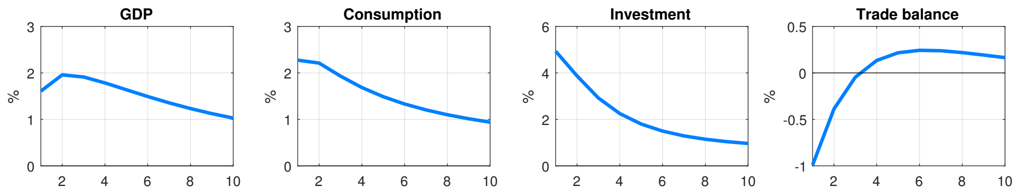
(b) Commodity sector productivity shock $\epsilon_t^{\tilde{a}}$



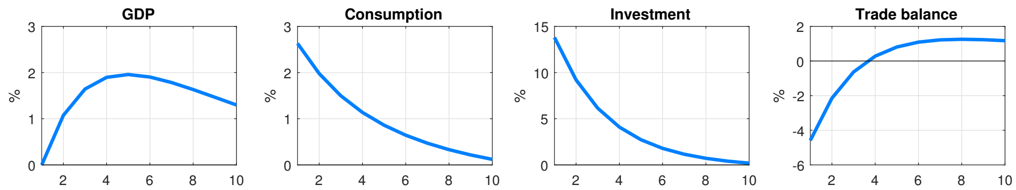
(c) Growth shock ϵ_t^g



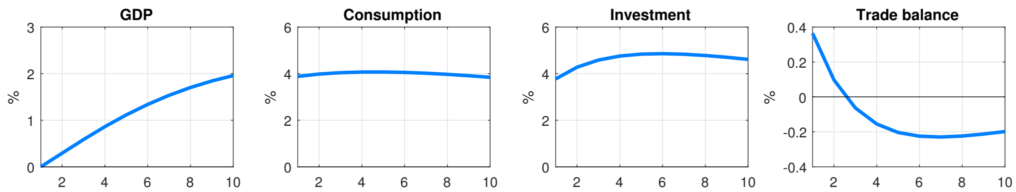
(d) Commodity price shock $\epsilon_t^{\tilde{p}}$



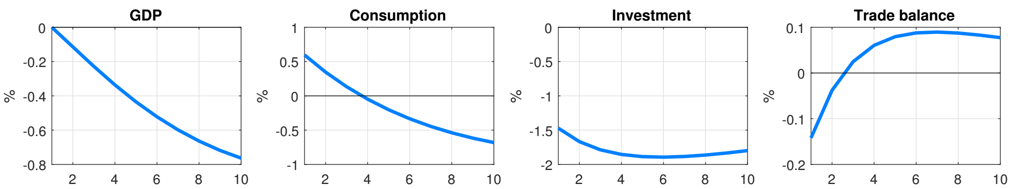
(e) Interest rate shock ϵ_t^{μ}



(f) Spending shock ϵ_t^s

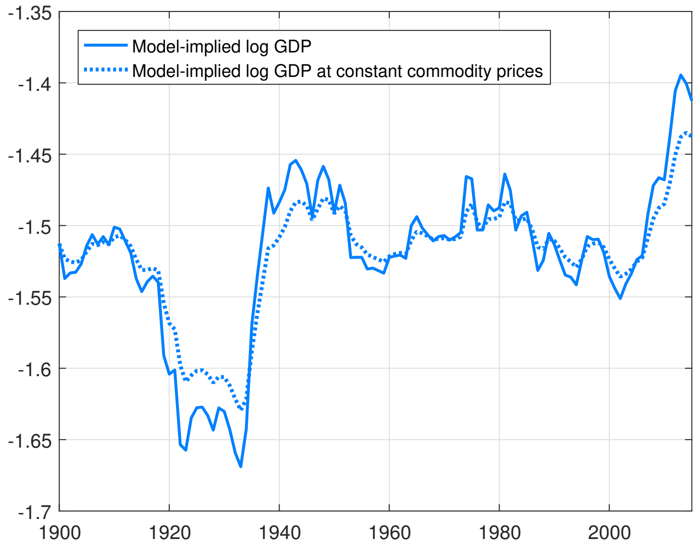


(g) Preference shock ϵ_t^{ν}



Note: All shocks have been re-scaled to give the same maximum GDP growth response as the commodity price shock in the body of the paper.

Figure 2.10: SIMULATED GDP UNDER DIFFERENT PRICE MEASUREMENT



Note: The blue solid line shows the economy's GDP, computed as in (2.19), when feeding in commodity prices and holding all other disturbances constant. The dotted line repeats the same exercise but computes GDP at the steady state relative price of commodities, that is, $\tilde{p}_t = \tilde{p}$.

Table 2.8: POSTERIOR ESTIMATES OF PARAMETERS

Parameter	Prior mean	Posterior mean	90% HPD interval	
ξ	0.199	0.2212	0.1550	0.2876
ψ	2.8	3.2057	2.5050	3.8984
ρ_a	0.5	0.8277	0.7494	0.9092
$\rho_{\tilde{a}}$	0.5	0.5887	0.2827	0.8980
ρ_g	0.5	0.5244	0.3199	0.7299
ρ_ν	0.5	0.8687	0.8382	0.8996
ρ_s	0.5	0.6440	0.5075	0.7832
ρ_μ	0.5	0.9199	0.8743	0.9693
$\rho^1_{\tilde{p}}$	0.8	0.8060	0.6840	0.9388
$-\rho^2_{\tilde{p}}$	0.15	0.1278	0.0105	0.2298
σ_a	0.10	0.0295	0.0231	0.0360
$\sigma_{\tilde{a}}$	0.10	0.0525	0.0242	0.0810
σ_g	0.10	0.0261	0.0193	0.0327
σ_ν	0.10	0.4582	0.4145	0.5000
σ_s	0.10	0.1876	0.1659	0.2089
σ_μ	0.10	0.0547	0.0410	0.0683
$\sigma_{\tilde{p}}$	0.10	0.1765	0.0876	0.2652

2.7.5 Interest rate premia and commodity prices: Simple formal illustration

Suppose there is a borrower who borrows amount D_t. With probability λ she is able to repay in full. With probability $1-\lambda$ only a repayment smaller than the borrowed amount D_t can be made. This repayment is a fraction ϕ of commodity output $\tilde{p}_t\tilde{y}_t$ (equivalently, $\tilde{p}_t\tilde{y}_t$ can be thought of as collateral which the lender can seize when full repayment is not possible). The presence of a risk-neutral lender who herself can obtain funds at the risk-free rate r^* and who faces perfect competition, will result in the following zero profit condition:

$$(1+r^*)D_t = \lambda(1+r_t)D_t + (1-\lambda)\phi\tilde{p}_t\tilde{y}_t, \tag{2.49}$$

which can be rearranged to

$$r_t = \frac{1+r^*}{\lambda} - \frac{1-\lambda}{\lambda D_t}\phi\tilde{p}_t\tilde{y}_t - 1. \tag{2.50}$$

As can be seen from (2.50), an increase in $\tilde{p}_t$ reduces the interest rate r_t, ceteris paribus. This is the key assumption of our model we aim to rationalize with the above illustration. Furthermore, and also consistent with our formulation in (2.9), r_t is increasing in the level of debt D_t.

2.8 Bibliography

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

Akinci, O. (2017): "Financial frictions and macroeconomic fluctuations in emerging economies," *Working paper.*

Antolin-Diaz, J., T. Drechsel, and I. Petrella (2017): "Tracking the Slowdown in Long-Run GDP Growth," *The Review of Economics and Statistics*, 99, 343–356.

Bastourre, D., J. Carrera, J. Ibarlucia, and M. Sardi (2012): "Common drivers in emerging market spreads and commodity prices," Tech. rep., Working Paper, Central Bank of Argentina (BCRA).

Ben Zeev, N., E. Pappa, and A. Vicondoa (2017): "Emerging economies business cycles: The role of commodity terms of trade news," *Journal of International Economics*, 108, 368 – 376.

Bernanke, B. S., M. Gertler, and S. Gilchrist (1999): "The financial accelerator in a quantitative business cycle framework," *Handbook of Macroeconomics*, 1, 1341 – 1393.

Cavallo, A. (2013): "Online and official price indexes: Measuring Argentina's inflation," *Journal of Monetary Economics*, 60, 152–165.

Cespedes, L. F. and A. Velasco (2014): "Was this time different?: Fiscal policy in commodity republics," *Journal of Development Economics*, 106, 92–106.

Chang, R. and A. Fernandez (2013): "On the sources of aggregate fluctuations in emerging economies," *International Economic Review*, 54, 1265–1293.

Farias, L. and M. E. A. da Silva (2017): "Commodity Prices and Business Cycles in Small Open Economies: The Role of News Shocks," *Working paper.*

Fernández, A., A. González, and D. Rodriguez (2015): *Sharing a Ride on the Commodities Roller Coaster*, IDB Working Paper.

Fernández, A., S. Schmitt-Grohé, and M. Uribe (2017): "World Shocks, World Prices, And Business Cycles An Empirical Investigation," *Journal of International Economics*, 108.

Ferreres, O. (2005): "Dos siglos de economía argentina," *Fundación Norte y Sur.*

Frankel, J. (2011): "How can commodity producers make fiscal and monetary policy less procyclical?" *Beyond the Curse: Policies to Harness the Power of Natural Resources, R.Arezki, T.Gylfason, A.Sy, eds. (IMF, 2011).*

Frankel, J. A. (2004): "Experience of and Lessons from Exchange Rate Regimes in Emerging Economies," *In Monetary and Financial Integration in East Asia: The Way Ahead, edited by Asian Development Bank, (Palgrave Macmillan).*

Gali, J. and T. Monacelli (2005): "Monetary policy and exchange rate volatility in a small open economy," *The Review of Economic Studies*, 72, 707–734.

García-Cicco, J., R. Pancrazi, and M. Uribe (2010): "Real Business Cycles in Emerging Countries?" *American Economic Review*, 100, 2510–31.

Gavin, M., R. Hausmann, and L. Leiderman (1996): "Macroeconomics of capital flows to latin america: Experience and policy issues," *Inter-American Development Bank, Research Department Working Papers Number 4012*.

Greenwood, J., Z. Hercowitz, and G. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78, 402–17.

Grilli, E. R. and M. C. Yang (1988): "Primary commodity prices, manufactured goods prices, and the terms of trade of developing countries: what the long run shows," *The World Bank Economic Review*, 2, 1–47.

Guimaraes, B. (2011): "Sovereign default: Which shocks matter?" *Review of Economic Dynamics*, 14, 553 – 576.

Kaminsky, G. L., C. M. Reinhart, and C. A. Vegh (2005): "When It Rains, It Pours: Procyclical Capital Flows and Macroeconomic Policies," *NBER Macroeconomics Annual 2004*, 19, 11–53.

Kehoe, T. J. and K. J. Ruhl (2008): "Are shocks to the terms of trade shocks to productivity?" *Review of Economic Dynamics*, 11, 804 – 819.

Kose, M. A. (2002): "Explaining business cycles in small open economies: 'How much do world prices matter?'," *Journal of International Economics*, 56, 299–327.

Lubik, T. A. (2007): "Non-stationarity and instability in small open-economy models even when they are closed," *Economic Quarterly*, 393–412.

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

——— (1995): "The terms of trade, the real exchange rate, and economic fluctuations," *International Economic Review*, 101–137.

Min, H.-G., D.-H. Lee, C. Nam, M.-C. Park, and S.-H. Nam (2003): "Determinants of emerging-market bond spreads: Cross-country evidence," *Global Finance Journal*, 14, 271 – 286.

Mitchener, K. J. and G. Pina (2016): "Pegxit Pressure: Evidence from the Classical Gold Standard," Working Paper 22844, National Bureau of Economic Research.

Neumeyer, P. A. and F. Perri (2005): "Business cycles in emerging economies: the role of interest rates," *Journal of Monetary Economics*, 52, 345–380.

Pfaffenzeller, S., P. Newbold, and A. Rayner (2007): "A short note on updating the Grilli and Yang commodity price index," *The World Bank Economic Review*, 21, 151–163.

Prasad, E., K. Rogoff, S.-J. Wei, and M. A. Kose (2006): "Effects of financial globalization on developing countries: some empirical evidence," *Ann Harrison, editor, Globalization and Poverty (University of Chicago Press: Chicago). Occasional Paper No. 220, International Monetary Fund, 2003. NBER WP 10942, 2004.*

Reinhart, C. and V. Reinhart (2009): "Capital Flow Bonanzas: An Encompassing View of the Past and Present," in *NBER International Seminar on Macroeconomics 2008*, National Bureau of Economic Research, Inc, NBER Chapters, 9–62.

Schmitt-Grohe, S. and M. Uribe (2003): "Closing small open economy models," *Journal of International Economics*, 61, 163 – 185.

——— (2017): "How Important Are Terms Of Trade Shocks?" *International Economic Review, Forthcoming.*

Shousha, S. (2016): "Macroeconomic Effects of Commodity Booms and Busts: The Role of Financial Frictions," *Unpublished Manuscript.*

Sims, C. A. and T. Zha (1999): "Error Bands for Impulse Responses," *Econometrica*, 67, 1113–1155.

Smets, F. and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

Taylor, A. M. (2014): "The Argentina Paradox: Microexplanations and Macropuzzles," Working Paper 19924, National Bureau of Economic Research.

Uribe, M. and V. Yue (2006): "Country spreads and emerging countries: Who drives whom?" *Journal of International Economics*, 69, 6–36.

Chapter 3

Agnostic Structural Disturbances (ASDs): Detecting and Reducing Misspecification in Empirical Macroeconomic Models

3.1 Introduction

Exogenous random shocks are the lifeblood of modern macroeconomic business cycle models. These shocks enter the model as innovations to structural disturbances that affect key aspects of the model. Whereas the prototype real business cycle (RBC) model features total factor productivity (TFP) as the only structural disturbance, recent generations of business cycle models typically include a multitude of structural disturbances. To avoid singularities when estimating a model, one needs at least as many unobserved random disturbances as observables. These random shocks can take the form of measurement error or structural disturbances. The larger the role for measurement error, the smaller the role of the theoretical model.[1] Thus, if researchers want the theoretical model to explain an important part of the data and they do not want to set aside information contained in additional observables, then they need to come up with a sufficiently large set of structural disturbances.

Incorporating structural disturbances correctly is nontrivial and it is not enough to have the right set. Structural disturbances impose (cross-equation) restrictions on model equations and, thus, on the model's solutions. Thus, each structural disturbance has to enter *each* model equation correctly. This is a real concern, since we often do not have independent evidence on how structural disturbances should affect the system. For example, should a risk-premium disturbance affect all Euler equations or only those of a specific type of investment? Is it correct to assume that structural disturbances are uncorrelated as is commonly done? Chari et al. (2007)

[1]There is a fundamental difference between measurement error and structural random disturbances. The latter are part of the economic model and their shocks affect the system through time according to the equations of the model. Measurement error does not.

propose "wedges" as alternatives to standard structural disturbances. However, it is important to realize that wedges also impose restrictions. For example, suppose one adds a "labor wedge" to the labor first-order condition. The assumption that this wedge does not enter any other equation implies restrictions on how this disturbance affects policy rules.

The contributions of this paper are threefold. First, based on a series of econometric estimation exercises using data generated from a known economic model, we document that a *minor* misspecification of the empirical model regarding structural disturbances can easily lead to *large* distortions for parameter estimates and model properties, such as business cycle statistics and impulse response functions (IRFs). Specifically, we consider the case where the empirical model wrongly excludes one of the structural disturbances and wrongly includes another. Everything else is correctly specified, including functional forms. Even though we adjust parameter values to ensure that this is a relatively minor type of misspecification, the results can be very drastic. For example, standard deviations as implied by the misspecified model are frequently multiples of their true values and correlation coefficients and IRFs can flip sign. These results are due solely to misspecification, since we use large samples and a consistent estimator.

Second, we propose agnostic structural disturbances (ASDs) as an alternative *structural* disturbance. In contrast to regular structural disturbances, ASDs impose *no* additional restrictions on policy rules. Nevertheless, they are very different from measurement error, because they are *structural* disturbances and propagate through the system like regular structural disturbances. Our ASD procedure can be used in two ways. First, it can be used to test whether regular structural disturbances are correctly specified. Second, an empirical specification can be enriched by adding ASDs as additional structural disturbances. Using Monte Carlo experiments, we document that the ASD procedure is capable of detecting and correcting for misspecification in samples of typical size.

The third contribution of our paper is to test whether the structural disturbances of the model in Smets and Wouters (2007) (SW) are correctly specified using the same US postwar data set. We find that the risk-premium and the investment-specific productivity disturbance are not correctly specified. We use our procedure to improve on the SW empirical specification. Specifically, our preferred specification (based on marginal likelihood considerations) has three ASDs and excludes the SW risk-premium and the SW investment-specific disturbance.

A nice feature of our procedure is that its outcomes provide insights into the nature of the agnostic disturbances. That is, although the ASD procedure itself does not rely on any theory, the estimation results – both the associated coefficients and their IRFs – may reveal a lot about the type of structural disturbance the data has identified.

One of the ASDs in our adjusted empirical specification of the SW model has a strong impact on the investment Euler equation and plays an important role for the fluctuations in investment. While the same is true for the standard investment-specific disturbance used in SW, our ASD enters the capital accumulation with a different sign than the investment-specific disturbance and also has a direct positive effect on capacity utilization. This ASD could capture an "investment-modernization" disturbance that positively affects the return on new investment, but goes together with an increased depreciation of *existing* capital. The latter would imply that this disturbance affects the capital accumulation with the opposite sign as a standard investment disturbance, consistent with our empirical results. The direct effect on utilization could compensate for this scrapping of existing vintages of capital.

The second ASD shares similarities with the SW risk-premium disturbance. Specifically, it plays a key role in the bond Euler equation. However, the way it enters the capital valuation equation indicates it is a preference disturbance, not a risk-premium disturbance. Interestingly, Smets and Wouters (2007) prefer the risk-premium disturbance over the preference disturbance of Smets and Wouters (2003) because it generates a positive comovement of the main economic aggregates, whereas a preference disturbance does not. Our ASD generates a typical business cycle even though it affects the capital valuation equation like a preference disturbance. The reason is that it also has an important impact on the investment Euler equation. Another noteworthy feature of this ASD is that it directly affects the policy rate. This indicates that the central bank responds differently to economic developments when these are due to changes in investors' required rates of return.

The third ASD has an important impact on the wage mark-up. Whereas the SW wage mark-up disturbance only affects one equation, our ASD also has an important effect on the capital value, the utilization, and the capital accumulation equation. Specifically, the data indicate that increased upward wage pressure goes together with more efficient use of capital. This disturbance has a very temporary impact. In contrast to the first two ASDs, this ASD does not replace a SW disturbance. Leaving the SW wage mark-up disturbance out of our preferred empirical specification reduces the marginal data density substantially. However, including this ASD does substantially lower the value of the MA coefficient in the ARMA representation of the SW wage mark-up disturbance.

In our application, we could give a sensible interpretation to each of the three ASDs like one often can do for wedges. Similar to wedges, there may not be a unique interpretation.[2] However, with wedges the researcher has to take a stand on where the wedges enter the model. The whole idea about our procedure is that it starts by being agnostic and it lets the data decide where and how ASDs should enter each model equation.

In the next section, we discuss the outcomes of our misspecification experiments, in which we generate data using the SW model as the data generating process (*dgp*)

[2] For example, regarding our third ASD, it is possible that it captures a higher wage mark-up that induces a more efficient use of capital. But it is also possible that it captures a desire to use capital more efficiently and that the higher wage mark-up is the price firms have to pay to obtain this efficiency increase.

and then estimate parameters with slightly misspecified empirical models. We use large samples, so the results are only due to misspecification and not to sampling variation. Section 3.3 provides a general discussion and motivation of our proposed misspecification detection and correction procedure. Section 3.4 describes how to use ASDs in practice. Section 3.5 documents the ability of ASDs to detect and correct for misspecification using Monte Carlo experiments for a typical application. We use again the SW model to generate data and the same type of misspecification of the empirical model as in section 3.2. But now we use a sample length of typical size. Section 3.6 discusses the results when our procedure is applied to the SW model on US data.

3.2 Large sample consequences of misspecification

In this section, we consider the consequences of estimating a misspecified empirical model. Specifically, we generate data with a known structural business cycle model and then estimate parameters using a misspecified empirical specification. We focus on large sample properties and use a Maximum Likelihood (ML) estimator, which is consistent in this environment. Thus, the results presented are not due to sampling variation. We document that even a minor misspecification can lead to substantial distortions in parameter estimates. These distortions matter in the sense that they imply model properties that are quite different from the true ones. In fact, even implied model moments for variables that are used as observables in the estimation can deviate substantially from their data counterparts (which represent the truth given that we focus on large sample properties).

3.2.1 The true underlying model for our experiment

We use the New Keynesian model of Smets and Wouters (2007), the workhorse model of empirical business cycle analysis, as the basis of our analysis. Parameter values of the true data generating process are set equal to those of the SW posterior mode.[3] The list of parameters estimated and their interpretation is given in Table 3.1.[4]

[3]The only exception is the parameter ρ_{ga}, which captures the impact of the TFP structural disturbance on the government expenditures structural disturbance. We set this coefficient equal to zero in both the true *dgp* and in the empirical model. This implies that all structural disturbances are uncorrelated. This is a typical assumption and makes our misspecification experiment more transparent. As discussed below, the misspecification considered is related to the specification of the set of structural disturbances. If $\rho_{ga} \neq 0$, then we would have to make additional choices whenever the misspecification involves either the TFP or the government spending shock. We explored some alternative cases in which $\rho_{ga} \neq 0$ and found similar results.

[4]We follow SW and do not estimate the depreciation rate, δ, the steady-state wage mark-up, $\overline{\mu}$, the steady-state level of government expenditures, $\overline{g}$, the curvature in the Kimball goods-market aggregator, ε_p, and the curvature in the Kimball labor-market aggregator, ε_w. Since we use demeaned data, we also fix the trend growth rate, $\overline{\gamma}$, the parameter controlling steady state hours, $\bar{l}$, the parameter controlling steady state inflation, $\overline{\pi}$, and the discount factor, β.

Table 3.1: PARAMETER EXPLANATIONS

α	Capital share
σ_c	Inverse IES of consumption
Φ	Fixed cost in production
ϕ	Elasticity of adjustment cost function
λ	Degree of consumption habits
ξ_w	Degree of wage rigidity
σ_ℓ	Inverse IES of leisure
ξ_p	Degree of price rigidity
ι_w	Degree of indexation for wages
ι_p	Degree of indexation for prices
ψ	Elasticity of capital utilization adj. cost function
r_π	Taylor rule coefficient on inflation
ρ	Degree of interest rate smoothing in Taylor rule
r_y	Taylor rule coefficient on output gap
$r_{\Delta y}$	Taylor rule coefficient on change in output gap
ρ_j	Persistence of exogenous disturbance j
μ_j	MA coefficient of exogenous disturbance j
σ_j	Standard deviation of exogenous disturbance j
	$j \in \{a, b, g, I, r, p, w\}$

Note: The table reports the parameters of the SW model that are estimated and their interpretation. The list of exogenous disturbances is given in the text.

3.2.2 The specification of the empirical model

The original SW model has seven exogenous random variables. Those are a TFP disturbance, ε_t^a, a risk-premium disturbance, ε_t^b, a government spending disturbance, ε_t^g, an investment-specific disturbance, ε_t^i, a monetary policy disturbance, ε_t^r, a price mark-up disturbance, ε_t^p, and a wage mark-up disturbance, ε_t^w. We leave out one of these seven disturbances when generating data for our misspecification experiments. The empirical specification also leaves out one disturbance, *but* not the right one. This means we have $7 \times 6 = 42$ experiments. Everything else is always correctly specified, including functional forms, specification of the processes for the exogenous random variables, and the values of the parameters that are not estimated. The observables used in SW consists of employment, the federal funds rate, the inflation rate, GDP, consumption, investment, and the real wage rate. We exclude the real wage rate so we have the same number of observables as structural disturbances which is consistent with the empirical exercise in SW.

Is this a likely misspecification? We believe that this type of misspecification is likely to be important in practice even if one includes a large set of structural disturbances. The first reason is that having a large set does not necessarily imply one includes all the true disturbances. Moreover, one does not only need to include all true disturbances, *each* disturbance has to enter *each* model equation correctly. For example, a TFP disturbance is typically modeled as a labor-augmenting productivity shock, but productivity changes could affect the production function differently.

Moreover, TFP increases may also affect other aspects of the production process such as the depreciation rate.[5]

Is this a "minor" misspecification? When generating the data, we adjust the standard deviation of the disturbance that is incorrectly excluded from the empirical specification to ensure that it is responsible for at most 10% of the volatility for *any* of the six observables used in the estimation. By doing this we reduce the quantitative importance of the misspecification.

One could argue that a misspecification is only minor if one would not detect it in a typical data set using some model selection criterion such as the marginal likelihood. This is a very strict requirement. Comparing a misspecified model with the correct one requires that researchers are aware of the correct specification and test their empirical model against it. Since structural disturbances can enter models in many different ways, researchers may not consider the correct one even if they consider several alternatives.[6] In this section, we use very large samples and misspecified models would be rejected against the truth. In section 3.2.6, we select two of the forty-two experiments and data sets of typical length and document using a computer intensive Monte Carlo analysis that a test comparing the misspecified model with the correct model would often *not* lead to a rejection of the misspecified model. This supports our claim that the misspecification considered here is indeed minor.

Estimation procedure. DSGE models are typically estimated with Bayesian techniques, which means that the estimation outcome is a weighted combination of the prior and the empirical likelihood. Misspecification of the empirical model affects the latter. With a tight prior, observed data – and thus misspecification of the likelihood – matter less for posterior estimates. Then, the quality of those estimates will depend on the quality of the prior. This paper focuses on the question how misspecification affects what the observed data imply for parameter estimates and implied model properties. Thus, we focus on the likelihood and use Maximum Likelihood estimation.[7]

In practice, there could be interesting interactions between the misspecification of the empirical model and small sample properties of the estimator. We abstract

[5]Similarly, Cúrdia and Reis (2012) argue that assumptions about the correlation of structural disturbances are important and that one can question the standard assumption that structural disturbances in macroeconomic models are not correlated.

[6]Indeed, although the SW empirical specification is a very carefully constructed model that incorporates insights of many previous empirical studies, it is still rejected against some minor modifications, as is shown in 3.6.

[7]Our optimization problem is relatively well defined. It helps, of course, that our experiments rely on very large samples and on empirical models that are only misspecified in terms of the driving processes. Moreover, we use the true parameter values as the initial conditions for the optimization routine and we specify bounds for the parameter values. These choices decrease computing time and also give a misspecified model the best possible chance to deliver estimates that are close to the truth. The innovation standard deviations of the disturbances are restricted to be in the interval $[0, 10]$ and the coefficients of their time series process in the interval $[0, 99]$. Given our focus on misspecified disturbances, we want these intervals to be large. For the structural parameters we set the lower bound and the upper bound to the first and ninety-ninth percentile according to the SW prior, centered at the parameter values of the true *dgp*.

from small sampling variation by using a large enough sample. In particular, our experiments are based on a sample of 10,000 observations. Our estimator is consistent and estimates are very close to the truth when the empirical model is correctly specified. Section 3.5 studies the small sample properties in detail for two out of the forty-two misspecification experiments.

Priors on the standard deviation of structural disturbances typically do not allow for point mass at zero. Ferroni et al. (2015) point out that this biases the results towards a positive role of all structural disturbances.[8] This is not an issue for us, since we use ML estimation. In fact, estimated standard deviations of disturbances that are part of the empirical model but *not* part of the true *dgp* turn out to be often close to zero.[9]

Identification. In appendix 3.7.2, we document that the estimated parameter values are identified using a strong version of the identification test of Komunjer and Ng (2011). This is true according to the correct and the misspecified empirical model. Thus, none of the results should be driven by non-identification rather than misspecification. Further justification for this claim is given in section 3.2.5.

3.2.3 Misspecification: Consequences for parameter values

Table 3.2 reports some key percentiles (across experiments) to characterize the range of the estimated parameter values. We only consider parameters that are in both the true and empirical specification.[10] All parameters are affected by misspecification to some extent. Moreover, the minor misspecifications considered in these forty-two experiments lead to massive distortions for several parameter estimates.

The median parameter estimates (across experiments) are relatively close to the true parameter values. Thus, our choice of experiments does not favor bias in a particular direction. There is one exception. The median value of the estimated standard deviation of the productivity disturbance innovation, σ_a, is equal to 0.92 compared to a true value of 0.45. The reason is that this disturbance often “absorbs" the variation of the disturbance that is not included in the empirical specification. Thus, the disturbance that is wrongly included in the empirical specification does not necessarily fulfill this role.

[8] There are several differences between their and our setup. They only consider one specific misspecified empirical model whereas we consider forty-two. Although they consider a limited Monte Carlo experiment (with 100 replications), the main discussion focuses on particular sample of 200 observations. In this section, we abstract from small sample issues by focusing on one very long sample. Most importantly, their main focus is on the consequences of using an inverse gamma prior for parameters that could well be zero. Our focus is on the misspecification of the empirical model, not the specification of the prior.

[9] In those cases, the role of the structural disturbance that is wrongly excluded from the empirical specification is “taken over" by some of the correctly included disturbances, not the one that is wrongly included.

[10] Specifically, for the parameters of the exogenous random processes, the experiments in which the disturbance is part of the empirical model – but not part of the true *dgp* – are excluded from the calculations of the percentiles.

Table 3.2: Parameter estimates across misspecification experiments

	Truth	Imposed Min	Min	10%	25%	Median	75%	90%	Max	Imposed Max
α	0.19	0.07	0.07	0.11	0.17	0.19	0.20	0.23	0.31	0.31
σ_c	1.39	0.53	0.53	0.78	1.14	1.35	1.60	1.82	2.25	2.25
Φ	1.61	1.33	1.33	1.33	1.53	1.77	1.89	1.89	1.89	1.89
ϕ	5.48	1.99	2.71	3.59	5.47	7.38	8.97	8.97	8.97	8.97
λ	0.71	0.45	0.45	0.59	0.71	0.74	0.84	0.89	0.90	0.90
ξ_w	0.73	0.47	0.50	0.67	0.73	0.75	0.82	0.87	0.91	0.92
σ_ℓ	1.92	0.18	0.18	0.18	0.52	1.87	2.71	3.66	3.66	3.66
ξ_p	0.65	0.40	0.53	0.60	0.65	0.78	0.86	0.86	0.86	0.86
ι_w	0.59	0.24	0.24	0.27	0.38	0.58	0.61	0.80	0.89	0.89
ι_p	0.22	0.01	0.01	0.01	0.10	0.22	0.32	0.48	0.63	0.65
ψ	0.54	0.20	0.20	0.20	0.42	0.54	0.68	0.86	0.86	0.86
r_π	2.03	1.45	1.45	1.45	1.71	2.07	2.39	2.61	2.61	2.61
ρ	0.81	0.53	0.62	0.73	0.79	0.81	0.85	0.88	0.92	0.97
r_y	0.08	-0.04	-0.04	0.01	0.05	0.09	0.16	0.20	0.20	0.20
$r_{\Delta y}$	0.22	0.10	0.10	0.10	0.10	0.20	0.24	0.34	0.34	0.34
ρ_a	0.95	0.00	0.50	0.82	0.92	0.96	0.98	0.99	0.99	0.99
ρ_b	0.18	0.00	0.04	0.09	0.13	0.17	0.26	0.36	0.80	0.99
ρ_g	0.97	0.00	0.94	0.96	0.97	0.97	0.99	0.99	0.99	0.99
ρ_I	0.71	0.00	0.57	0.60	0.68	0.71	0.78	0.84	0.95	0.99
ρ_r	0.12	0.00	0.01	0.06	0.11	0.13	0.18	0.33	0.50	0.99
ρ_p	0.90	0.00	0.70	0.77	0.84	0.89	0.93	0.96	0.98	0.99
ρ_w	0.97	0.00	0.93	0.95	0.97	0.97	0.98	0.99	0.99	0.99
μ_p	0.74	0.00	0.08	0.22	0.43	0.73	0.82	0.91	0.95	0.99
μ_w	0.88	0.00	0.00	0.00	0.87	0.89	0.92	0.96	0.98	0.99
σ_a	0.45	0.00	0.42	0.47	0.67	0.92	1.49	2.57	3.20	10
σ_b	0.24	0.00	0.07	0.20	0.23	0.24	0.26	0.27	0.29	10
σ_g	0.52	0.00	0.52	0.52	0.52	0.53	0.55	0.56	0.57	10
σ_I	0.45	0.00	0.14	0.25	0.39	0.44	0.46	0.48	0.54	10
σ_r	0.24	0.00	0.22	0.23	0.23	0.24	0.26	0.28	0.31	10
σ_p	0.14	0.00	0.04	0.09	0.12	0.14	0.15	0.16	0.17	10
σ_w	0.24	0.00	0.18	0.20	0.21	0.24	0.25	0.29	0.31	10

Note: This table gives information about the parameter estimates across the forty-two misspecification experiments. For the parameters of the laws of motion of the disturbances, we exclude an experiment from the calculations of the percentiles when the disturbance is part of the empirical model, but not part of the true *dgp*. The table also reports the bounds imposed on parameter estimates. See Table 3.1 for the definitions of the parameters.

Even if we exclude cases for which the estimates fall in the bottom or top 10%, then we find that estimates are substantially different from their true value for many parameters. For example, for the labor supply elasticity with respect to the real wage, σ_l, the 10^{th} percentile is equal to 0.18 and the 90^{th} percentile is equal to 3.66, compared with a true value of 1.92. For the parameter capturing the indexation of wages ι_w, the same two percentiles are 0.27 and 0.8, compared with a true value of 0.59. For the parameter capturing the indexation of prices, ι_p, the two numbers are 0.01 and 0.48, compared with a true value of 0.22. When the two 10% tails are not excluded and the full range of estimates is considered, then the range substantially increases. Specifically, the largest values are 0.89 and 0.63 for the indexation of wages and prices, respectively.[11] Recall that these distortions are solely due to misspecification, not to small-sample variation.

For several parameters, the results remain bad when we narrow the range of outcomes considered. For example, when we exclude the bottom and the top 25%, then the values for σ_l, vary between 0.52 and 2.71 compared with a true value of 1.92. The results are also quite bad for ϕ, the elasticity in the capital adjustment cost function, for which the 25^{th} percentile is equal to 5.47 and the 75^{th} percentile is equal to 8.97.

3.2.4 Missspecification: Consequences for model properties

The previous section documents that misspecification can lead to large distortions in parameter values. Parameter estimates are often of interest in themselves. At least as important are the properties of the estimated structural model. It could be that different parameter configurations lead to similar model properties. In this section, we address this by looking at implied moments and IRFs.

Implied model moments

We begin by documenting the consequences of model misspecification for implied model moments using the misspecification setup described above. Table 3.3 reports the range of values for typical business cycle properties as implied by the estimated parameter values of the forty-two experiments considered. Specifically, it reports standard deviations and correlation coefficients relative to their true values. Thus, a value equal to 1 means that there is no distortion. The column labeled "true value" reports the range of values the corresponding moment has according to the true *dgp*.[12]

Misspecification implies an upward bias for volatility in our experiments.[13] This

[11] Parameter estimates are constrained to be in a range, and the largest estimate of the wage indexation parameter is constrained by the imposed upper bound.

[12] Moments are not the same across experiments, since we adjust the standard deviations of the structural disturbances to ensure that the wrongly omitted disturbance does not play an important role.

[13] Section 3.2.3 documents an upward bias for σ_a, the standard deviation of the TFP disturbance. Since one disturbance is missing from the empirical model, it is not surprising that there is a shift towards some of the other disturbances. By contrast, here we find an upward bias for *total*

Table 3.3: Moments: Ratio of implied value to truth across experiments

	True value	Min	10%	25%	Median	75%	90%	Max
				(estimates, scaled by true value)				
Std(y_t)	[3.48 , 5.12]	0.51	0.78	0.92	1.03	1.64	4.46	6.03
Std(c_t)	[3.30 , 5.58]	0.45	0.76	0.92	1.03	1.81	4.12	6.62
Std(i_t)	[9.73 , 12.94]	0.70	0.87	0.99	1.11	1.71	3.81	6.47
Std(r_t)	[0.52 , 0.61]	0.76	0.90	0.94	1.00	1.36	2.28	2.78
Std(π_t)	[0.37 , 0.54]	0.64	0.72	0.94	1.01	1.25	2.21	2.98
Std(w_t)	[2.13 , 2.70]	0.73	0.83	0.92	1.08	2.28	5.57	10.87
Corr(y_t, c_t)	[0.65 , 0.94]	0.28	0.68	0.93	0.99	1.07	1.15	1.52
Corr(y_t, i_t)	[0.74 , 0.87]	0.69	0.83	0.95	1.00	1.10	1.16	1.29
Corr(c_t, i_t)	[0.63 , 0.89]	-0.68	0.60	0.92	1.00	1.19	1.34	1.57
Corr(c_t, r_t)	[-0.65 , -0.35]	-0.71	0.54	0.86	0.99	1.11	1.52	2.13
Corr(i_t, w_t)	[0.29 , 0.69]	-1.52	0.10	0.64	1.07	1.49	1.99	3.28
Corr(i_t, π_w)	[0.51 , 0.80]	0.36	0.84	0.97	1.02	1.17	1.34	1.75

Note: This table reports the outcomes across experiments for the indicated moment as implied by parameter estimates relative to its true value. Thus a value equal to 1 indicates that there is no distortion due to misspecification. Each row reports percentiles across our forty-two experiments. It also reports the range of values of the true moments across the experiments. All moments considered are related to variables that are used in the estimation as observables.

upward bias could be specific to our particular type of misspecification. However, the observed upward bias is consistent with the simple analytical example discussed in appendix 3.7.1.[14] The results are solely due to misspecification, since we use very large samples and our ML estimator is consistent when the empirical model is correctly specified.

The overestimation of volatility is enormous in some cases. Even if we exclude the top 25%, then standard deviations can be multiples of the true standard deviation. For example, the 75$^{\text{th}}$ percentile for the standard deviation of wages is 2.28 times its true value. This ratio increases to 5.57 when we only exclude the top 10%. The 90$^{\text{th}}$ percentiles for the consumption and output standard deviation ratios are 4.12 and 4.46, which also indicates massive over-prediction. The 90$^{\text{th}}$ percentile number for investment is equal to 3.81 and in the worst experiment the implied standard deviation is 6.47 times as big as the true value. By contrast, the values in the lower tail are less drastic. Excluding the bottom 10%, we find that the largest distortions are found for inflation for which the 10$^{\text{th}}$ percentile is 0.72, that is, implied volatility is 28% below its true value. If we consider all experiments, then the smallest ratio is equal to 0.45, which is found for the implied standard deviation of consumption.

Misspecification also has large quantitative implications for correlation coefficients. In fact, the sign of the correlation coefficient as implied by parameter estimates turns out to be different from its sample analogue in several cases. This would not be a big deal if the two correlation coefficients are both close to zero. But there

variability.

[14]In appendix 3.7.1, we discuss a simple example which documents analytically how maximum likelihood estimation of a misspecified model can lead to an *arbitrarily* large upward bias in the implied variance of an observable.

are also cases in which the implied correlation coefficient according to the estimated empirical model and the true correlation coefficient are both large in absolute value and differ in sign.[15,16]

Impulse response functions (IRFs)

To conclude the discussion on the consequences of misspecification, we document that misspecification can also have a large impact on impulse response functions. There are many IRFs to consider. Figure 3.1 plots for three IRFs the outcomes across the experiments and documents that the distortions can be large. We exclude the cases when the disturbance of interest is in the empirical specification, but not part of the true *dgp*. It would not be surprising if these are different.[17] Thus, the disturbance of interest is part of the true *dgp* as well as the empirical model for all three cases considered.

Figure 3.1a plots the response of output to a TFP disturbance. This is obviously a key characteristic of the model. The black line plots the true IRF and the grey lines plot the IRFs as implied by the empirical model for the different experiments. All IRFs are based the same size shock.[18] If the grey lines are close to the black line, then misspecification of the empirical model has only minor consequences for the IRF considered. The sign of the IRF is virtually always correct and TFP disturbances always have a noticeable positive impact on aggregate output.[19] Nevertheless, the figure documents that there are large differences in terms of initial impact, overall magnitude, shape, and persistence.

Figure 3.1b plots the response of the real wage to a monetary policy shock. This is clearly the kind of model property one would want to get right when analyzing monetary policy. The figure shows again a wide variety of responses across the different empirical specifications. Whereas the true response is substantial, there are several empirical specifications that predict a very small change. There are also a few specifications that give a much larger response. We want to reemphasize that the plotted IRFs are for a disturbance that is correctly included in the empirical model.

Figure 3.1c reports the results for the inflation IRF of an investment-specific shock. For most experiments the IRFs display a similar pattern, but there are important differences in terms of magnitude. For three experiments, however, the IRFs are completely at odds with the true IRF. Whereas the true IRF is positive and has reverted back to zero after twenty periods, the IRFs implied by these three misspeci-

[15] A striking example is the experiment in which the government disturbance is not present in the true *dgp* and the empirical model excludes the risk-premium disturbance instead. The true correlation between consumption and investment is equal to 0.67 whereas the one implied by the estimated model is equal to -0.41.

[16] The smallest correlation coefficient (in absolute value) according to the true model is 0.29, so any sign change implies a nontrivial change in the correlation coefficient.

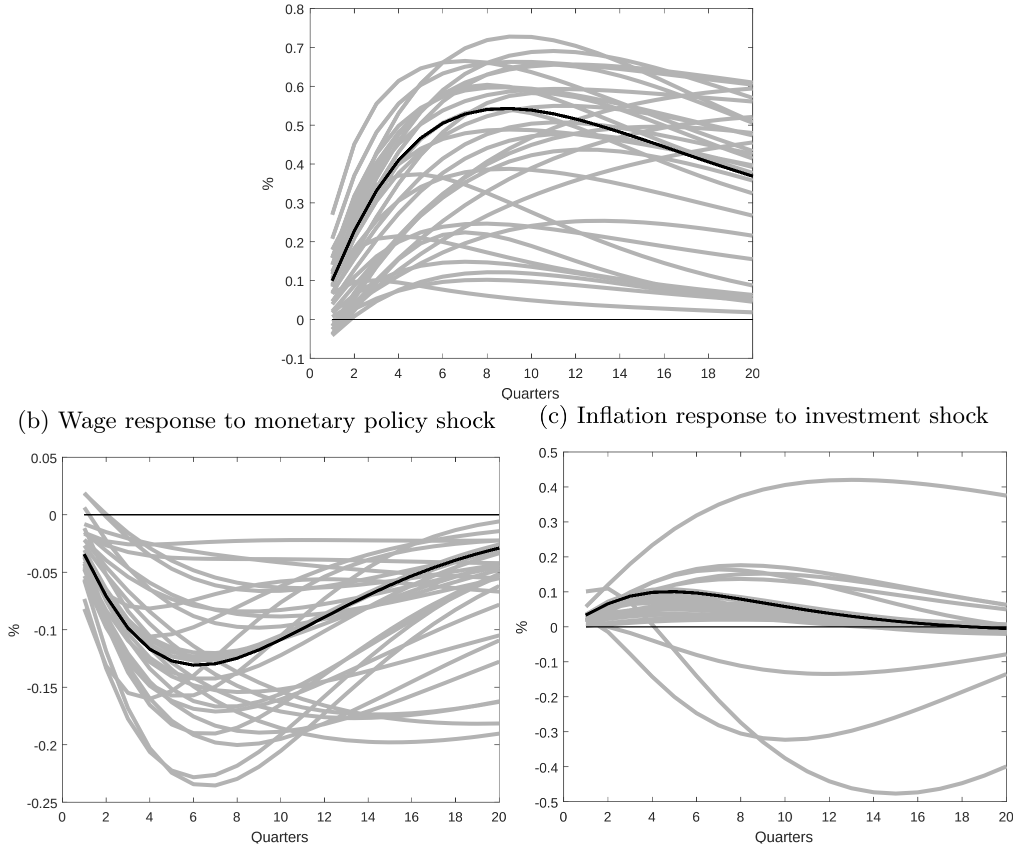
[17] Also, we cannot calculate IRFs for a particular disturbance if that disturbance is not part of the empirical specification. This means that each figure plots IRFs for thirty-two cases.

[18] That is, one standard deviation according to the original SW model. Differences across IRFs are bigger if we use the estimated standard deviations for the different experiments.

[19] In some experiments, the initial response is negative. However, its value is then very small.

Figure 3.1: IRFs according to true and misspecified empirical models

(a) Output response to TFP shock



Note: The figure plots the true IRF (black)and the IRFs implied by the misspecified (grey) empirical models considered. The results are based on a very large sample, so results are not due to small sample variation. These IRFs are for shocks that are correctly included in the model. Also, we do not use estimated standard deviations, but use the same size shock for all IRFs.

fied empirical models are negative and indicate larger volatility and more persistence. Again, relatively small changes in parameter values can change these IRFs such that they are much closer to the true IRF.[20]

3.2.5 Is weak identification the cause?

In appendix 3.7.2, we demonstrate that all parameters are identified in all models considered.[21] Moreover, we use a very large sample to estimate the parameters so the large range of values for parameter estimates cannot be caused by samples being too short to be informative. Also, the finding that the different parameter values are

[20]Specifically, if σ_c, the parameter controlling curvature in the utility function and λ, the parameter indicating the habit component in the utility function, are set equal to their true values, then these three IRFs have a shape that is similar to the true IRF, that is, also predict a hump-shaped positive response. The responses still differ somewhat from the truth in having a more delayed response and a more persistent effect. The estimated values for σ_c in the three experiments are 0.65, 0.53, and 0.53, whereas the true value if equal to 1.39. The estimated values for λ are equal to 0.86, 0.87, and 0.85, whereas the true value is equal to 0.71.

[21]All true specifications have one structural disturbance less than the original SW model. This turns out not to matter for identification. In fact, estimated parameters remain identified when we do the identification test for specifications with five disturbances that exclude the disturbance that is not part of the true *dgp* as well as the one that is erroneously omitted from the empirical specification.

associated with quite different model properties indicates that the results discussed in this section are not due to parameters not being identified. As a final check, we compare the values of the likelihood according to the misspecified model at the estimated values and the true values. When using the true values, we do re-estimate the parameters of the exogenous random variables.[22] The smallest difference between the two log likelihood values is equal to 14.5 and there are only four experiments for which the difference is less than 100. The mean (median) difference is equal to $10,371$ (5501).[23]

3.2.6 Is the misspecification really minor?

The misspecification experiments considered above involve the inclusion of one wrong and the exclusion of one correct structural disturbance. Everything else is correctly specified. So the misspecification affects only a small part of all the model features researchers have to specify when writing down a complete empirical model.

Nevertheless, one could argue, that this misspecification is not that likely for the analysis in Smets and Wouters (2007), since SW was preceded by years of empirical analysis by many authors. However, in section 3.6, we document that we clearly reject the null that two of the included structural disturbances are correctly specified against several alternatives. It is important to recall that correct specification of a structural disturbance is not only getting the nature of the disturbance right, but also that it enters each model equation correctly. In section 3.6, we will argue that this is not the case for two of the seven SW structural disturbances.

Furthermore, it could be argued that a misspecification is only minor if the misspecified model is not rejected when its fit is compared with the fit of the correct model. This is, of course, a test that one could never implement in practice, since it requires knowing the truth. The large differences in likelihood discussed in section 3.2.5 indicate that the misspecified model would be easily rejected. However, those likelihood values correspond to tests using unrealistically large samples. The appropriate question is whether one would reject the misspecified model with a sample of typical length and typical estimation procedure.

To address this question, we do a Monte Carlo experiment in which the model is estimated as in SW. That is, the data set has the same number of observations, the parameters are estimated with the same Bayesian methodology, and the priors are also the same. We assess model fit using the marginal data density (MDD). These are expensive Monte Carlo experiments.[24] Therefore, we consider only two of the possible

[22]This is a conservative choice, since differences in the likelihoods would be larger if these parameters are not re-estimated.

[23]It is not surprising that across experiments, there are some for which the misspecification is smaller than for others resulting in smaller differences between the two likelihood values. After all, our experiments are not designed to find large misspecification. Our set is constructed using a simple variation in the set of the original structural disturbances.

[24]The reason is that they involve an optimization problem containing many parameter values. In contrast to the exercise in section 2, the optimization here is a bit more difficult, since it is affected by small-sample sampling variation. Moreover, it has to be repeated for every Monte Carlo replication.

forty-two misspecification experiments of section 3.2. They were chosen as follows. We ranked all experiments by the likelihood value of the misspecified specification relative to the likelihood of the correct specification. The idea is that misspecification is less severe if the likelihood values are close to each other. The first experiment chosen is the one corresponding to the sixty-sixth percentile and the second is the one corresponding to the thirty-third percentile.[25] Thus, our experiments are neither the least nor the most problematic in terms of misspecification. In section 3.5, we return to these two examples and we will document that consequences of misspecification are severe for both cases.

For the experiment at the thirty-third percentile we find that the misspecified model has a higher marginal data density in 17.8% of the Monte Carlo replications. Thus, one would prefer the wrong empirical model over the correct one in about four out of five cases in cases *if* one is so lucky to be able to do the test against the true model specification.

For the experiment at the sixty-sixth percentile, this number decreases to 52%. That is, the correct and the misspecified model have roughly an equal chance of having the best fit when realistic samples are used.[26]

3.3 Agnostic Structural Disturbances

In this section, we develop and motivate our “structural agnostic disturbance" (ASD) procedure to detect and correct for misspecification. ASDs can be added to a structural model and they can be used to test whether a regular structural disturbance is correctly specified.

3.3.1 Underlying theoretical model

Consider the following linearized model

$$0_{n\times 1} = \mathbb{E}_t \left[\Lambda_2 (\Psi) s_{t+1} + \Lambda_1 (\Psi) s_t + \Lambda_0 (\Psi) s_{t-1} + \Gamma (\Psi) \varepsilon_{t+1} + \Upsilon (\Psi) \varepsilon_t\right], \quad (3.1a)$$

$$\varepsilon_t = G\varepsilon_{t-1} + H\eta_t, \quad (3.1b)$$

$$\mathbb{E}_t \left[\eta_{t+1}\right] = 0, \quad (3.1c)$$

$$\mathbb{E}_t \left[\eta_{t+1}\eta'_{t+1}\right] = I_{m\times m}, \quad (3.1d)$$

where Ψ is the vector containing the structural parameters, s_t is the $n \times 1$ vector of endogenous variables, and ε_t is the $m \times 1$ vector of exogenous random variables.

[25] The first (second) Monte Carlo experiment corresponds to the case when the true *dgp* does not include a monetary policy (TFP) disturbance, but the empirical model leaves out the investment disturbance instead.

[26] This Monte Carlo experiment does indicate an interesting aspect of sampling variation. The large sample analysis indicates that the empirical model considered in this second Monte Carlo experiment is more misspecified than the one of the first Monte Carlo, since there was a large difference in the marginal likelihoods of the correct and misspecified model. In terms of marginal likelihoods, this ranking is reversed in small samples.

All variables are defined relative to their steady state values. Most linearized DSGE models can be represented with such a system of equations.[27]

Type of misspecification considered. As in section 3.2, the misspecification focuses on the modeling of the structural disturbances. That is, whether the included disturbances are the right ones and whether the restrictions they impose on the model equations correct.

3.3.2 The ASD procedure

There are two ways to describe and implement the ASD procedure. The first formulation is discussed in section 3.3.2. This formulation highlights that our procedure is more general than the procedure that adds wedges to particular model equations. We provide the second formulation in section 3.3.2 after discussing some background information in section 3.3.2. This second formulation makes clear that our procedure is more efficient than the misspecification procedures that combine a DSGE model with a reduced-form empirical model as in Ireland (2004) and Del Negro et al. (2007). This efficiency advantage is made possible by focusing on one particular type of misspecification, namely exogenous disturbances not being the right ones or not being modeled correctly. As explained below, this allows us to use some of the structure of the model.

We will show that these two formulations are not different procedures, but different ways to implement this procedure.[28] Which procedure is more convenient in practice will depend on the application.

ASD procedure: First formulation based on model equations

Consider the model given in equation (3.1). To simplify the exposition, we start with the case for which s_t includes only state variables and all n state variables are observables. Suppose that the researcher is only sure about m_1 structural disturbances. These are part of the vector, $\varepsilon_{1,t}$. If $m_1 < n$ and there are no other disturbances, then there is a singularity problem. One option would be to add measurement error. But structural disturbances and measurement errors are very different. Structural disturbances affect economic variables and propagate through the system according to the economic mechanisms of the model. Measurement error disturbances do not.[29] Another option is to make a best guess and to add a vector $\varepsilon_{2,t}$ with m_2 additional

[27]Linearization leads to accurate solutions for many business cycle models. When this is not the case, then this is an additional source of misspecification.

[28]See section 3.3.2.

[29]See section 3.3.2 for an explanation. Moreover, most researchers would find it undesirable if "measurement" error explains a large part of the data.

structural disturbances with $m_2 \geq n - m_1$. Equation (3.1) can then be written as

$$
\begin{aligned}
0_{n\times 1} &= \mathbb{E}_t \left[\Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} + \Gamma(\Psi)\varepsilon_{t+1} + \Upsilon(\Psi)\varepsilon_t\right] \\
&= \mathbb{E}_t \left[\begin{array}{c} \Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} \\ + \left[\Gamma_1(\Psi)\ \Gamma_2(\Psi)\right] \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix} + \left[\Upsilon_1(\Psi)\ \Upsilon_2(\Psi)\right] \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \end{array} \right], \qquad (3.2a)
\end{aligned}
$$

$$
\begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1} \\ \varepsilon_{2,t-1} \end{bmatrix} + \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \end{bmatrix}, \tag{3.2b}
$$

$$
\mathbb{E}_t \begin{bmatrix} \eta_{1,t+1} \\ \eta_{2,t+1} \end{bmatrix} = 0, \tag{3.2c}
$$

$$
\mathbb{E}_t \left[\begin{bmatrix} \eta_{1,t+1} \\ \eta_{2,t+1} \end{bmatrix} \begin{bmatrix} \eta_{1,t+1} & \eta_{2,t+1} \end{bmatrix} \right] = I_{m\times m}. \tag{3.2d}
$$

The column vectors $\Gamma_2(\Psi)$ and $\Upsilon_2(\Psi)$ capture the restrictions imposed by the m_2 additional structural disturbances. In the remainder of this section, we document that no such restrictions are imposed when agnostic structural disturbances are added.

Adding ASDs to model equations. If one adds agnostic structural disturbances instead of regular structural disturbances, then the system of equations is modified as follows:

$$
0_{n\times 1} = \mathbb{E}_t \left[\begin{array}{c} \Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} \\ + \left[\Gamma_1(\Psi)\ \mathbf{\Gamma}_2\right] \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix} + \left[\Upsilon_1(\Psi)\ \mathbf{\Upsilon}_2\right] \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \end{array} \right]. \tag{3.3}
$$

The key aspect of our procedure is that $\mathbf{\Gamma}_2$ and $\mathbf{\Upsilon}_2$ are reduced-form coefficients that do *not* contain any restrictions on Ψ. Moreover, when $G_{21} = 0$, which is typically the case as structural disturbances are usually modeled to be uncorrelated, then

$$
\mathbb{E}_t \left[\mathbf{\Gamma}_2 \varepsilon_{2,t+1} + \mathbf{\Upsilon}_2 \varepsilon_{2,t}\right] = \left(\mathbf{\Gamma}_2 G_{22} + \mathbf{\Upsilon}_2\right) \varepsilon_{2,t}. \tag{3.4}
$$

Using this insight, we can write the system as

$$
0_{n\times 1} = \mathbb{E}_t \left[\begin{array}{c} \Lambda_2(\Psi) s_{t+1} + \Lambda_1(\Psi) s_t + \Lambda_0(\Psi) s_{t-1} \\ + \Gamma_1(\Psi) \varepsilon_{1,t+1} + \Upsilon_1(\Psi) \varepsilon_{1,t} + \widehat{\mathbf{\Upsilon}}_2 \varepsilon_{2,t} \end{array} \right], \tag{3.5}
$$

where $\widehat{\mathbf{\Upsilon}}_2 = G_{22}\mathbf{\Gamma}_2 + \mathbf{\Upsilon}_2$. All that matters for the model is $\widehat{\mathbf{\Upsilon}}_2$, which means that adding an agnostic disturbance introduces one additional parameter for each model equation.[30,31] Replacing regular structural disturbances with agnostic structural disturbances may make it harder to identify Ψ, the structural parameters of the model.

[30] Without loss of generality one can set the standard deviations of the innovation of the ASDs equal to 1, which in this case is a normalization of the diagonal elements of $H_{2,2}$. As with regular structural disturbances, one would need to estimate the parameters of the time series specification contained in G.

[31] As discussed later, one could choose to leave the agnostic disturbance out of some equations.

As discussed in appendix 3.7.2, this turned out to be not an issue for the experiments discussed in this paper. Identification of $\widehat{\boldsymbol{\Upsilon}}_2$ will be discussed in section 3.3.2.

Useful proposition for second ASD formulation

In this section, we will proof a proposition that will be helpful with the second formulation of the ASD procedure. Consider again the model given in equation (3.2), which divides the vector with exogenous disturbances, ε_t, into two parts, the $m_1 \times 1$ vector, $\varepsilon_{1,t}$, and the $m_2 \times 1$ vector, $\varepsilon_{2,t}$. A recursive solution to equation (3.2) has the following form:

$$\begin{aligned} s_t &= A(\Psi)\, s_{t-1} + B(\Psi)\varepsilon_t \\ &= A(\Psi)\, s_{t-1} + \begin{bmatrix} B_1(\Psi) & B_2(\Psi) \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}. \end{aligned} \tag{3.6}$$

The following proposition states that the properties of $\varepsilon_{2,t}$ do *not* affect the coefficients of the policy rule related to s_{t-1} and $\varepsilon_{1,t}$, that is, they do not affect $A(\Psi)$ and $B_1(\Psi)$. Thus, it does not matter whether $\varepsilon_{2,t}$ is a regular or an agnostic structural disturbances and the time series properties of $\varepsilon_{2,t}$ do not matter either. The only assumption needed is that the elements of G_{21} are equal to zero, which corresponds to the case when $\varepsilon_{1,t}$ has no effect on *future* values of $\varepsilon_{2,t}$. This is not very restrictive given that the literature usually sets all elements of G_{21} equal to zero (and also all elements of G_{12}, $H_{1,2}$, and $H_{2,2}$ as well as the off-diagonal elements of G_{11}, G_{22}, $H_{1,1}$ and $H_{2,2}$).

Proposition 1. *If the model is given by equation* (3.2) *and all elements of* G_{21} *are equal to zero, then* (*i*) $A(\Psi)$ *and* $B_1(\Psi)$ *do not depend on* $\Gamma_2(\Psi)$ *and* $\Upsilon_2(\Psi)$, *which characterize the nature of the additional disturbances, and* (*ii*) $A(\Psi)$ *and* $B_1(\Psi)$ *do not depend on* G_{22}, H_{21}, *and* H_{22}, *which characterize the time series properties of* $\varepsilon_{2,t}$.

Proof. Substitution of the policy rule as given in equation (3.6) into the system of equations (3.2) gives,

$$0_{n\times 1} = \left(\Lambda_2 A^2 + \Lambda_1 A + \Lambda_0\right) s_{t-1} + \left(\Lambda_2 AB + \Lambda_2 BG + \Lambda_1 B + \Gamma G + \Upsilon\right) \varepsilon_t, \tag{3.7}$$

$$\varepsilon_t = \begin{bmatrix} \varepsilon_{1,t} & \varepsilon_{2,t} \end{bmatrix}', \tag{3.8}$$

$$B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}, \tag{3.9}$$

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}, \tag{3.10}$$

where we have suppressed the dependence of coefficients on Ψ. The first equation has to hold for all values of s_{t-1} and ε_t. This implies that a solution must satisfy

$$\Lambda_2 A^2 + \Lambda_1 A + \Lambda_0 = 0_{n\times n} \tag{3.11}$$

and

$$\Lambda_2 AB + \Lambda_2 BG + \Lambda_1 B + \Gamma G + \Upsilon = 0_{n\times(m_1+m_2)}. \tag{3.12}$$

A does not depend on the time series properties of $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, since B, G, and H do not appear in equation (3.11). Equation (3.12) can be written as follows

$$\overline{\Lambda}\begin{bmatrix} B_1 & B_2 \end{bmatrix} + \Lambda_2 \begin{bmatrix} B_1 & B_2 \end{bmatrix}\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} + \Gamma \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} + \Upsilon = 0_{n\times(m_1+m_2)}, \tag{3.13}$$

where $\overline{\Lambda} = \Lambda_2 A + \Lambda_1$. This is a system of $n \times (m_1 + m_2)$ equations to solve for the elements of B. It can be split into the following two sets of systems:

$$\begin{aligned} \overline{\Lambda}B_1 + \Lambda_2 B_1 G_{11} + \Lambda_2 B_2 G_{21} + \Gamma_1 \begin{bmatrix} G_{11} \\ G_{21} \end{bmatrix} + \Upsilon_1 &= 0_{n\times m_1}, & (3.14) \\ \overline{\Lambda}B_2 + \Lambda_2 B_1 G_{12} + \Lambda_2 B_2 G_{22} + \Gamma_2 \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix} + \Upsilon_2 &= 0_{n\times m_2}. & (3.15) \end{aligned}$$

If $G_{21} = 0$, then equation (3.14) contains $n \times m_1$ equations to solve for all the elements of B_1. The solution cannot depend on G_{22} or H_2 since these matrices do not appear in this equation.■

It is intuitive that the elements of G_{21} have to be equal to zero, that is, $\varepsilon_{1,t}$ should not affect future values of $\varepsilon_{2,t}$. If current values of $\varepsilon_{1,t}$ do affect future values of $\varepsilon_{2,t}$ and therefore future values of s_t, then one has to know how $\varepsilon_{2,t}$ affects model outcomes to determine how $\varepsilon_{1,t}$ affects current outcomes for s_t.

ASD procedure: Second formulation based on policy functions

The second formulation highlights the differences with alternative procedures that deal with misspecification by combining a DSGE model and a VAR. This alternative formulation is also useful in terms of understanding whether adding agnostic disturbances leads to identification issues.

An alternative way of writing the solution to the model is the following:

$$\begin{aligned} s_t &= \sum_{i=1}^{m} s_t^{[i]}, & (3.16) \\ s_t^{[i]} &= A(\Psi)\, s_{t-1}^{[i]} + B_{\cdot,i}(\Psi)\, \varepsilon_{i,t}, & (3.17) \end{aligned}$$

where $s_t^{[i]}$ represents the outcome of the state variable if the *only* disturbance in the economy is the i^{th}-disturbance, $\varepsilon_{i,t}$, and $B_{\cdot,i}$ is the i^{th} column of B. Thus, one can think of the s_t variables as the sum of the outcomes in "one-disturbance" economies. The linearity of the model is important for this additive property. According to proposition 1, the coefficients on the lagged state variable, $A(\Psi)$, do not depend on the particular disturbance considered. That is, whereas $B_{\cdot,i}(\Psi)$ is indexed by i because it depends on what kind of disturbance is the driving force of the economy,

$A(\Psi)$ does not. This property greatly increases the efficiency of our procedure.

Our proposed procedure consists of including m_2 agnostic structural disturbances. This results in the following time series representation of the policy functions:[32]

$$
\begin{aligned}
s_t &= \sum_{i=1}^{m} s_t^{[i]}, && (3.18a)\\
s_t^{[i]} &= A(\Psi)\, s_{t-1}^{[i]} + B_{\cdot,i}(\Psi)\, \varepsilon_{i,t} \quad \text{for } i \le m_1, && (3.18b)\\
s_t^{[i]} &= A(\Psi)\, s_{t-1}^{[i]} + \boldsymbol{B}_{\cdot,i} \varepsilon_{i,t} \qquad \text{for } m_1+1 \le i \le m_1+m_2 = m. && (3.18c)
\end{aligned}
$$

In terms of notation, $B_{\cdot,i}(\Psi)$ contains coefficients associated with a regular structural disturbance which are a function of Ψ and bold font $\boldsymbol{B}_{\cdot,i}$ contains reduced-form coefficients associated with a structural agnostic disturbance. The *only* difference between this specification and the standard DSGE specification with only regular structural disturbances is that the $\boldsymbol{B}_{\cdot,i}$ coefficients are unrestricted reduced-form coefficients. Since our agnostic disturbances are structural disturbances, their impact propagates through the system exactly as regular structural disturbances do, that is, as described by $A(\Psi)$. The property of linear models that $A(\Psi)$ does not depend at all on what is the nature of the structural disturbances nor on their time series properties makes it possible to efficiently add structural disturbances to the specification without having to be specific on what they are.

The dimension of $\boldsymbol{B}_{\cdot,i}$ is equal to n, the number of state variables. This means that adding an agnostic disturbance means estimating an additional n parameters. The number of additional parameters to be estimated is limited because structural disturbances differ in their initial impact, but their propagation through time is the same for all disturbances and controlled by $A(\Psi)$. Moreover, an increase in the standard deviation of an agnostic structural disturbance affects the model variables in exactly the same way as an identical proportional increase of the elements of $\boldsymbol{B}_{\cdot,i}$. Consequently, the standard deviation of an agnostic disturbance can be normalized to equal 1.[33] If there are observables that are not state variables, then one also needs an additional equation for these y_t variables, which for our set-up is given by

$$
\begin{aligned}
y_t &= \sum_{i=1}^{m} y_t^{[i]}, && (3.19a)\\
y_t^{[i]} &= C(\Psi)\, s_{t-1}^{[i]} + D_{\cdot,i}(\Psi)\, \varepsilon_{i,t} \quad \text{for } i \le m_1, && (3.19b)\\
y_t^{[i]} &= C(\Psi)\, s_{t-1}^{[i]} + \boldsymbol{D}_{\cdot,i} \varepsilon_{i,t} \qquad \text{for } m_1+1 \le i \le m_1+m_2 = m, && (3.19c)
\end{aligned}
$$

where y_t is the $(\overline{n} \times 1)$ vector with observables that are not state variables. Each additional observable used in the estimation will introduce one more coefficient related

[32] According to proposition 1, this specification is valid as long as the elements of G_{12} are equal to zero, which is usually the case.

[33] If the time series processes of the two disturbances have the same number of parameters, then replacing a regular structural disturbance by an agnostic disturbance typically means estimating an additional $n-1$ parameters. The number would be less if some structural parameters are associated only with the regular structural disturbance that is replaced.

to the agnostic structural disturbances.

Identification

Replacing $B_{\cdot,i}(\Psi)$ with $\boldsymbol{B}_{\cdot,i}$ reduces the number of restrictions on structural parameters, which could affect the identification of Ψ. We have verified that the structural parameters, Ψ, continue to satisfy the local identification conditions as specified in Komunjer and Ng (2011) when we replace regular structural disturbances by ASDs. The coefficients of $\boldsymbol{B}_{\cdot,i}$ are also identified locally since they directly enter the policy functions. However, there is no global identification of the $\boldsymbol{B}_{\cdot,i}$ coefficients when $m_2 > 1$, since the agnostic disturbances are interchangeable, that is, there is no difference between say the first and the second agnostic disturbance in how they affect model equations. This is a consequence of being agnostic.

The first formulation of our procedure adds agnostic disturbances to the model equations. Under what conditions are the associated coefficients, i.e., the elements of $\widehat{\boldsymbol{\Upsilon}}_2$ identified? What matters for identification are the policy functions, that is, the $\boldsymbol{B}_{\cdot,i}$ and the $\boldsymbol{D}_{\cdot,i}$ coefficients. These coefficients are a function of the $\widehat{\boldsymbol{\Upsilon}}_2$ coefficients. Since the total number of coefficients in $\boldsymbol{B}_{\cdot,i}$ and $\boldsymbol{D}_{\cdot,i}$ is equal to the number of state variables plus the number of observables that are not state variables, $n + \overline{n}$, one can add the agnostic disturbance to at most $n + \overline{n}$ model equations for each of the elements of $\widehat{\boldsymbol{\Upsilon}}_2$ to be identified. This is a necessary, not a sufficient condition.[34]

Identification of the elements of $\widehat{\boldsymbol{\Upsilon}}_2$ only becomes important if one wants to give an economic interpretation of the agnostic disturbance. As discussed in section 3.6, however, this can be a useful exercise.

Equivalence of first and second formulation

The easiest case to consider is the one in which the model consists of $n+\overline{n}$ equations and the observables are the n state variable plus $\overline{n}$ other observables, where $\overline{n}$ could be zero. If the agnostic disturbance is added to all $n + \overline{n}$ model equations, then the two different ways to implement the procedure are identical.

The first formulation, which adds agnostic disturbances to model equations, is more flexible. The reason is that it allows us to add the agnostic disturbances to only a subset of the $n + \overline{n}$ equations. By excluding the agnostic disturbance from some equations one does impose restrictions on the agnostic disturbance and this implies that the first and the second implementation will lead to different policy functions and different estimation results. Imposing such restrictions moves us away from being fully agnostic, but there may be cases where this flexibility of the first formulation is very useful. In section 3.6, we document how model selection procedures can be used to impose restrictions leading to more concise formulations that make it possible to interpret the agnostic disturbances.

[34]To understand why this is not a sufficient condition consider a system that consists of two equations containing the model's two state variables, $s_{1,t}$ and $s_{2,t}$, and no other variables. Also, y_t satisfies the equation $y_t = 2s_{1,t} + s_{2,t}$. If y_t is not an observable, then one could not add the agnostic disturbance to this equation, because its associated coefficient would, of course, not be identified.

Now consider the case when the model has more than $n + \overline{n}$ equations, that is, some model variables are not state variables or observables, and the agnostic disturbance is added to more than $n + \overline{n}$ model equations. From the discussion above, we know that not all the elements of $\widehat{\boldsymbol{\Upsilon}}_2$ can be identified. That is, different combinations of the coefficients in $\widehat{\boldsymbol{\Upsilon}}_2$ lead to the same values for the $n+\overline{n}$ coefficients in $\boldsymbol{B}_{\cdot,i}$ and $\boldsymbol{D}_{\cdot,i}$. As long as the agnostic disturbance remains agnostic and there is no need to interpret the $\widehat{\boldsymbol{\Upsilon}}_2$ coefficients, then this is not a problem. Specifically, it does not affect the identification of the structural parameters Ψ.[35]

Comparison with alternative procedures

In this section, we discuss how our procedure compares with alternatives proposed in the literature. A detailed description of these alternative approaches can be found in appendix 3.7.5.

Agnostic structural disturbances versus wedges. Equations (3.2) and (3.5) point out the difference between adding an agnostic structural disturbance and adding a regular structural disturbance. Adding a regular structural disturbance requires specifying in which equation the disturbance appears and how the associated elements of $\Gamma_2(\Psi)$ and $\Upsilon_2(\Psi)$ depend on the structural parameters Ψ. Adding an agnostic disturbance does not impose such restrictions. Wedges are similar to *regular* structural parameters in that they only appear in a subset of equations. Sometimes only one equation. Wedges may or may not impose restrictions on the structural parameters, Ψ. For example, one of the wedges considered in Chari et al. (2007) is a productivity disturbance. This disturbance appears in the budget constraint and the first-order condition for capital and imposes cross-equation parameter restrictions. By contrast, when a "labor wedge" is added to the labor-supply first-order condition, then this does not impose restrictions on the structural parameters, since it does not appear in any other equation. Relative to an agnostic disturbance, however, it is restrictive because it is not allowed to appear in other model equations.[36]

Agnostic structural disturbances versus measurement error. ASDs differ from measurement error in that the latter is not a *structural* disturbance. Consequently, its impact on the different elements of *model* variables does not propagate through the system as structural disturbances do. To understand this difference

[35] In practice, a good optimization routine should still be able to find the true optimized value of the objective function and associated values for Ψ even though it may take some time before it realizes that several variations in the elements of $\widehat{\boldsymbol{\Upsilon}}_2$ do not lead to improvements in the target.

[36] Inoue et al. (2015) provide a formal analysis for using wedges to detect and identify misspecification. Using a New Keynesian model, they introduce a labor wedge into the cost minimization problem of the intermediate good producing firm, and a final good wedge and a bond demand wedge into the household budget constraint. Similar to the productivity disturbance, such wedges only appear in a limited set of equations and do impose parameter restrictions.

consider the following system of equations:

$$\begin{aligned} s_t &= As_{t-1} + B\varepsilon_t, && (3.20) \\ y_t &= Cs_t + D\varepsilon_t, && (3.21) \\ \varepsilon_t &= G\varepsilon_{t-1} + H\eta_t. && (3.22) \end{aligned}$$

The first equation represents a very simple structural model that governs the law of motion of the state variable, s_t. The second equation specifies the relationship between the observable, y_t and the state variable. ε_t is a scalar exogenous random variable. A value of C equal to 1 means that the state variable is the observable. If ε_t is measurement error, then $D \neq 0$ *and* $B = 0$. That is, measurement error affects the difference between data and model variables, but does not affect how model variables behave. By contrast, if ε_t is a structural disturbance, then $D = 0$ and $B \neq 0$. Now ε_t does affect model variables and propagates through the system according to the structural model, that is, according to equation (3.20).

The idea of the agnostic procedure is to add not one but two different ASDs to equations (3.20) and (3.21). This procedure would allow for several possibilities discussed above *and* combinations, namely one or two structural disturbances with no measurement error, one structural disturbances that is correlated with measurement error, one structural disturbance and uncorrelated measurement error, or just measurement error.

Agnostic disturbances versus a DSGE-VAR. Ireland (2004) and Del Negro et al. (2007) combine a DSGE model with a reduced-form VAR that contains the observables. Specifically, they start with a fully specified DSGE model as represented by equations (3.18a), (3.18b), (3.19a), and (3.19b). Since they have no agnostic structural disturbances, the value of m_2 is equal to zero.

There are two key differences between these two approaches and ours. First, our approach focuses on a particular type of misspecification, which allows it to use aspects of the model that are not affected by this misspecification, namely $A(\Psi)$ and $C(\Psi)$. Second, introducing a VAR into the estimation means that the number of disturbances necessarily increases by a number equal to the number of variables in the VAR. Moreover, adding a VAR introduces many more parameters unless the number of observables is small. Our procedure allows for a more parsimonious approach and could consist of adding just one new disturbance or replacing one regular structural disturbance with an agnostic structural disturbance.

Both differences imply that our approach is more efficient in terms of the number of parameters that it has to be estimate.[37] The price of parsimony is that our procedure is not designed to detect misspecification unrelated to structural disturbances, that is, misspecification associated with restrictions imposed by $B(\Psi)$ and $D(\Psi)$.

[37] For example, for the popular DSGE model of Smets and Wouters (2007) with 7 observables, a VAR with 4 lags would mean estimating 204 additional coefficients. As discussed in section 3.6, the implementation of our procedure for this model means estimating twelve more parameters.

Although, it is not designed to do so, ASDs might very well pick up other types of misspecification such as wrong functional forms and time variation in structural parameters. The DSGE-VAR approach explicitly allows misspecification in $A(\Psi)$ and $C(\Psi)$. However, Chari et al. (2008) point out that the VAR with a finite number of lags that does not contain *all* the model's state variables is likely to be misspecified. This means that the DSGE-VAR approach cannot deal with all possible misspecifications either.

Another difference emerges as the sample size goes to infinity. With the DSGE-VAR approach one has two "competing" empirical specifications, a DSGE model and a VAR. Since every DSGE suffers from at least some minor misspecification, one can expect the VAR to fully take over as the sample size goes to infinity. If that happens, then one is left with a reduced-form model that can no longer be used for policy analysis. This will never happen with our approach, since the propagation of state variables will always be determined by $A(\Psi)$ and the relationship between state variables and observables by $C(\Psi)$. If the number of regular structural disturbances in the true data generating process is less than or equal to the number of agnostic structural disturbances, then one can expect the role of regular structural disturbances to be driven to zero as the sample size goes to infinity.[38] The restrictions imposed by $B(\Psi)$ and $D(\Psi)$ would then no longer play a role.

3.4 What to do in practice?

In this section, we first discuss how agnostic structural disturbances can be used as a test for misspecification. Next, we discuss how agnostic structural disturbances can be applied to reduce misspecification.

3.4.1 ASDs to test for misspecification

Agnostic structural disturbances differ from regular structural disturbances in that their *initial* impact on the economy is not restricted and, thus, imposes no restrictions on model parameters. As indicated in equation (3.5), regular structural disturbances, $\varepsilon_{1,t}$, enter model equations as $\Gamma_1(\Psi)\varepsilon_{1,t+1} + \Upsilon_1(\Psi)\varepsilon_{1,t}$, whereas agnostic structural disturbances, $\varepsilon_{2,t}$, enter models equations as $\widehat{\boldsymbol{\Upsilon}}_2\varepsilon_{2,t}$, where $\widehat{\boldsymbol{\Upsilon}}_2$ is a vector of reduced-form coefficients that does not impose restrictions on Ψ. Since these are two competing models, and the former is a restricted version of the latter, standard model selection statistics can be used to test whether the restrictions imposed by structural disturbances are correct.

Specifically, a simple and transparent way to proceed is to carry out a model selection test, such as a likelihood-ratio test, for each of the regular structural disturbance considered separately. For example, if the disturbance in question is a wage mark-up disturbance, then one first estimates the model with a wage mark-up disturbance and then re-estimates the model with the wage mark-up disturbance

[38] Assuming that there are enough ASDs to avoid any singularity issues.

replaced by an ASD. Let $\mathcal{L}(\Psi)$ be the log likelihood of the model with the wage mark-up disturbance and let $\mathcal{L}(\Psi, \widehat{\boldsymbol{\Upsilon}}_2)$, be the log likelihood of the model with the wage mark-up disturbance replaced by an ASD.[39] To test the restrictions imposed by the wage mark-up disturbance one checks whether $\mathcal{L}(\Psi, \widehat{\boldsymbol{\Upsilon}}_2) - \mathcal{L}(\Psi)$ exceeds the critical value of a $\chi^2(q)$ distribution with q degrees of freedom, where q is the difference in the number of parameters between the two models. One could compare marginal data densities if one prefers a Bayesian methodology.

One can also assess whether a particular regular structural disturbance is restrictive by looking at changes in parameter estimates and model properties after the regular structural disturbance has been replaced by an agnostic structural disturbance.

If the restrictions imposed by the wage mark-up disturbance are rejected, then one has two options. First, one could modify how the wage mark-up affects the model. In section 3.6, we show how the estimated model with agnostic disturbances provides useful insights for such modifications. Second, one could simply use the estimated model with the agnostic disturbance.

3.4.2 ASDs to reduce misspecification

To estimate models one would like to use all available observables. When estimating DSGE models one needs at least as many disturbances as observables to avoid a singularity problem. As the number of observables increases, it becomes more difficult to come up with sensible structural disturbances. Recall that it is not just a question of conjecturing a particular type of structural disturbance. The structural disturbance has to enter each and every model equation correctly. An alternative is to add ASDs.

Adding agnostic disturbances does not complicate the estimation in practice. For example, to add an agnostic disturbance, $\varepsilon_{2,t}$, to a model estimated with Dynare one would add $\widehat{\boldsymbol{\Upsilon}}_{i,2}\varepsilon_{2,t}$ to the i^{th} model equation, where $\widehat{\boldsymbol{\Upsilon}}_{i,2}$ is the i^{th} element of $\widehat{\boldsymbol{\Upsilon}}_2$. Under our second formulation, adding an ASD simply means adding an extra column to the policy rules with the ASD and its reduced-form coefficients.

3.5 ASDs and Misspecification: Small-Sample Monte Carlo experiments

In section 3.2, we considered the large-sample consequences of using a (slightly) misspecified empirical model which wrongly excluded one structural disturbance and included one that was not part of the true model. In that experiment, an empirical model that includes an ASD instead of the wrongly included regular structural

[39] These expressions are based on the notation of the first formulation of our ASD procedure. In terms of the notation of the second formulation, the log likelihood of the model with the agnostic disturbance would be denoted by $\mathcal{L}(\Psi, \boldsymbol{B})$ when all observables are state variables or by $\mathcal{L}(\Psi, \boldsymbol{B}, \boldsymbol{D})$ when some observables are not state variables.

disturbance, would uncover the true parameter values. The reason is that this ASD-augmented empirical model is correctly specified, the ML estimator is consistent, a large sample is used, and the structural parameters remain identified.[40]

In this section, we consider the same misspecification experiment, but now consider *small-sample* Monte Carlo experiments. This will allow us two answer two questions. First, is the ASD procedure effective in detecting misspecification when we compare the ASD-augmented empirical model with the misspecified empirical model? Second, what is the efficiency loss if one replaces a regular structural disturbance that is part of the true underlying model with an ASD? The empirical model remains correctly specified if one does so, but one looses efficiency because one estimates additional reduced-form parameters and imposes less true restrictions.

3.5.1 Experiments and empirical specifications.

As in section 3.2, the *dgp* is the SW model with six of the seven structural disturbances, but now we use a sample of typical length, namely 156, which is the same as the number of observations used to estimate the model in Smets and Wouters (2007). The number of Monte Carlo replications is equal to 1,000. In each Monte Carlo replication, we estimate model parameters using the SW empirical model that is identical to the true *dgp* and two additional specifications. The first is an empirical model, that is – as in section 3.2 – misspecified, because it excludes one of the structural disturbances of the true *dgp* and erroneously includes another. The second excludes the same structural disturbance, but now includes an ASD. This last specification is also correct. That is, there are values of the reduced-form coefficients such that the specification is identical to the original SW one.

These are computationally expensive Monte Carlo experiments. Therefore, we only consider two of the possible forty-two combinations to misspecify and those are the same two as those considered in section 3.2.6.

3.5.2 ASD misspecification test when alternative is misspecified

To evaluate whether the ASD procedure can detect misspecification, we first use a Likelihood Ratio (LR) test that compares the likelihood of the empirical specification with the agnostic disturbance to the likelihood of the misspecified empirical model. The number of degrees of freedom is equal to ten, since the agnostic specification has ten more parameters.[41] With this procedure, the ASD procedure rejects the misspecified model in all Monte Carlo replications in both experiments. The procedure is, thus, quite powerful in detecting misspecification. As discussed in section 3.2.6, however, if we use a Bayesian model comparison procedure based on SW priors, then the ASD procedure rejects the misspecified specification in 82.2% and 52% of the generated samples for the first and the second Monte Carlo experiment, respectively.

[40] In fact, parameters remain identified, if two regular structural disturbances are replaced by ASDs.

[41] We use the second formulation of our procedure. This formulation introduces the smallest possible number of additional parameters.

It is not surprising that the power reduces with a Bayesian approach. The reason is that the posteriors of the misspecified and the agnostic specification are more similar than their likelihood functions since the posteriors share the same prior.

3.5.3 ASD misspecification test when alternative is correct

Next, we do the same ASD test for misspecification when the alternative model is correctly specified. For the first Monte Carlo experiment, we find that the rejection rate is 21.5% at the 10%-level and 12% at the 5%-level. For the second experiment, these two numbers are 20.9% and 12.6%. Thus, the small-sample results do not coincide precisely with the theoretical predictions based on large-sample theory. However, the distortions are not that unreasonable. In appendix 3.7.3, we document that the histogram of estimated χ^2 statistics is reasonably close to the theoretical (large-sample) χ^2 distribution, but has a slightly fatter upper tail.

3.5.4 Correcting for misspecification

The discussion above made clear that the ASD procedure does very well in terms of detecting misspecified models and reasonably well in not rejecting correctly specified models in small samples. In this subsection, we document that the estimates of the structural parameters obtained with the agnostic procedure are much closer to the true values than those obtained with the misspecified empirical model. In fact, they are very similar to those obtained with the correctly specified fully-structural empirical model.

Table 3.4 reports the average absolute error of the parameter estimates relative to the true value for the three different empirical models across Monte Carlo replications. Consistent with the large-sample results discussed in section 3.2, parameter estimates obtained with the misspecified structural model are substantially worse than those obtained with the correctly specified model. The average of the errors for the misspecified model is more than twice as large as the one for the correctly specified model for several parameters.[42] Average errors for the misspecified model are typically better for the second experiment. However, that is not true for all parameters. For example, the average error for σ_c is substantially higher in the second experiment, whereas there is only a modest increase for the correctly specified model.

For the first Monte Carlo experiment, the average error outcomes for the agnostic setup and the correct specification are very similar. Although only slightly, the average error is actually lower for the agnostic specification for ten of the twenty-seven parameters. For the second Monte Carlo experiment, the fully specified SW specification comes with some noticeable efficiency advantages for several parameter estimates. Nevertheless, the estimates obtained using the agnostic procedure are still much better than the one obtained with the misspecified model.

[42]Particular problematic is the standard deviation of the TFP disturbance in the first Monte Carlo experiment for which the average error is almost nine time as large as the one for the correct empirical model. Consistent with the results of section 3.2, this disturbance often takes over the role of the wrongly excluded structural disturbance.

Figures 3.2 and 3.3 plot histograms characterizing the distribution of the parameter estimates across Monte Carlo replications for a selected set of parameters.[43] Each panel reports the results for the correctly specified model (dark line and dots), the agnostic procedure (white bars), and the misspecified model (blue/dark bars).

Table 3.4: Average absolute errors across Monte Carlo experiments

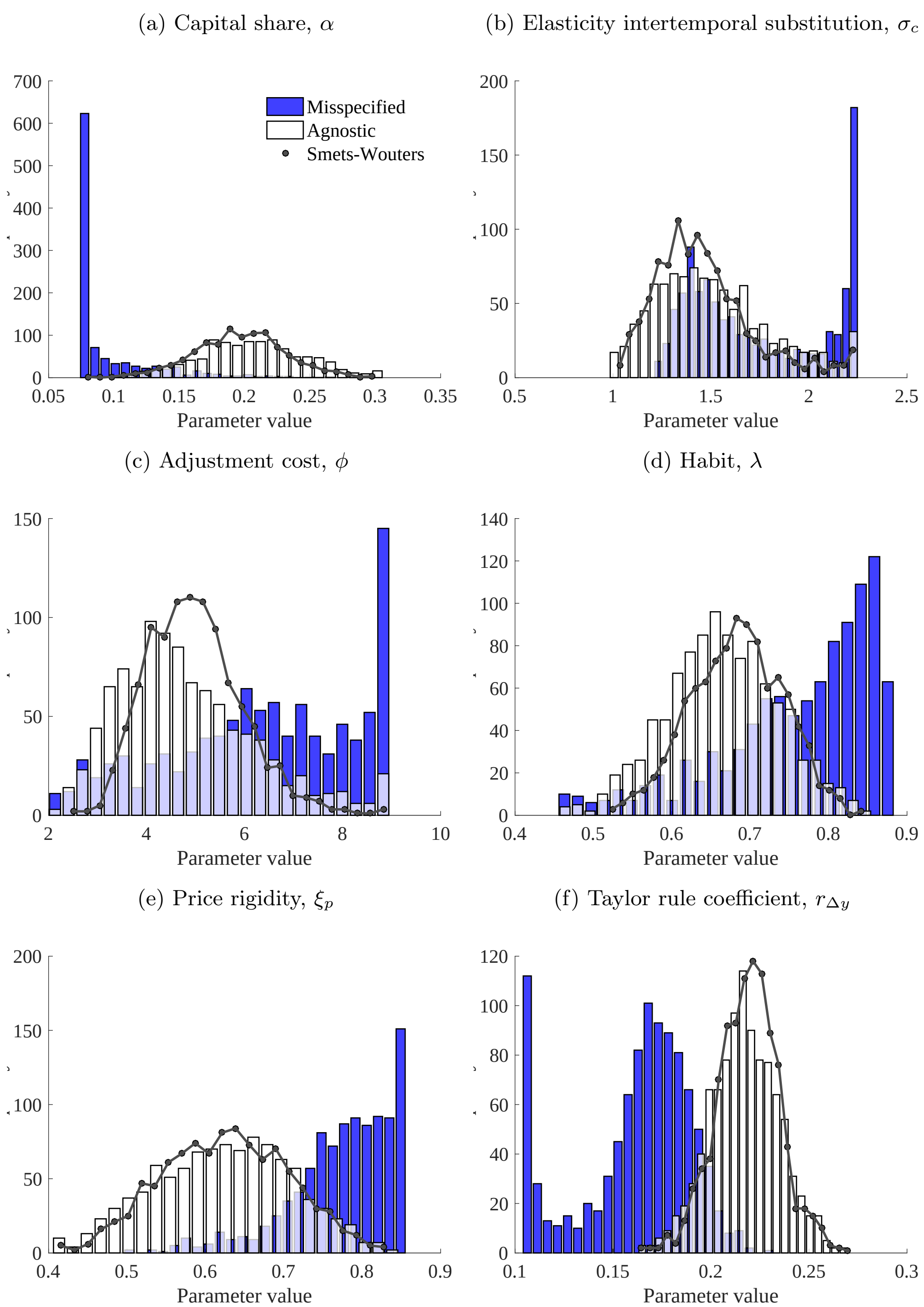
	true value	**average error first MC**			**average error second MC**		
		misspecified	agnostic	SW	misspecified	agnostic	SW
α	0.19	0.098	0.035	0.028	0.056	0.048	0.037
σ_c	1.39	0.384	0.246	0.191	0.540	0.288	0.226
Φ	1.61	0.217	0.212	0.191	0.192	0.212	0.164
ϕ	5.48	1.793	1.326	0.899	1.429	1.269	0.896
h	0.71	0.096	0.069	0.052	0.083	0.077	0.057
ξ_w	0.73	0.082	0.090	0.076	0.092	0.095	0.081
σ_ℓ	1.92	1.652	0.640	0.532	1.506	0.939	0.831
ξ_p	0.65	0.130	0.074	0.068	0.090	0.080	0.070
ι_w	0.59	0.205	0.165	0.159	0.190	0.168	0.160
ι_p	0.22	0.142	0.109	0.101	0.128	0.112	0.100
ψ	0.54	0.182	0.128	0.109	0.150	0.134	0.118
r_π	2.03	0.295	0.277	0.241	0.347	0.380	0.333
ρ	0.81	0.031	0.025	0.022	0.034	0.038	0.030
r_y	0.08	0.051	0.025	0.021	0.055	0.034	0.029
$r_{\Delta y}$	0.22	0.058	0.014	0.012	0.057	0.039	0.033
ρ_a	0.95	0.071	0.028	0.020	-	-	-
ρ_b	0.18	0.161	0.078	0.073	0.133	0.079	0.071
ρ_g	0.97	0.020	0.016	0.013	0.018	0.016	0.014
ρ_I	0.71	-	-	-	-	-	
ρ_r	0.12	-	-	-	0.089	0.072	0.067
ρ_p	0.90	0.181	0.090	0.067	0.188	0.070	0.053
ρ_w	0.97	0.031	0.030	0.019	0.022	0.029	0.021
μ_p	0.74	0.246	0.188	0.161	0.250	0.173	0.139
μ_w	0.88	0.071	0.072	0.056	0.069	0.071	0.057
σ_a	0.45	0.441	0.061	0.052	-	-	-
σ_b	0.24	0.050	0.021	0.021	0.040	0.023	0.021
σ_g	0.52	0.035	0.027	0.026	0.026	0.027	0.025
σ_I	0.45	-	-	-	-	-	
σ_r	0.24	-	-	-	0.013	0.015	0.014
σ_p	0.14	0.022	0.017	0.015	0.019	0.017	0.015
σ_w	0.24	0.026	0.021	0.020	0.022	0.023	0.021

Note: This table reports the average absolute error across Monte Carlo replications for the indicated parameter and empirical specification. See Table 3.1 for the definitions of the parameters. The first (second) Monte Carlo experiment corresponds to the case when the true *dgp* does not include the monetary policy (TFP) disturbance, but the empirical model leaves out the investment disturbance instead.

The figures document that the distributions of estimates obtained with the correct specification and the agnostic procedure are both qualitatively and quantitatively very similar. By contrast, the distribution of estimates obtained with the misspecified empirical model can be vastly different. For example, Panel (a) of Figure 3.2 documents that the distribution of estimates of the capital share parameter, α, displays a strong downward bias when the misspecified empirical model is used.

[43]A full set of results for all parameters is given in appendix 3.7.3.

Figure 3.2: Histograms: First Monte Carlo experiment



(a) Capital share, α

(b) Elasticity intertemporal substitution, σ_c

(c) Adjustment cost, ϕ

(d) Habit, λ

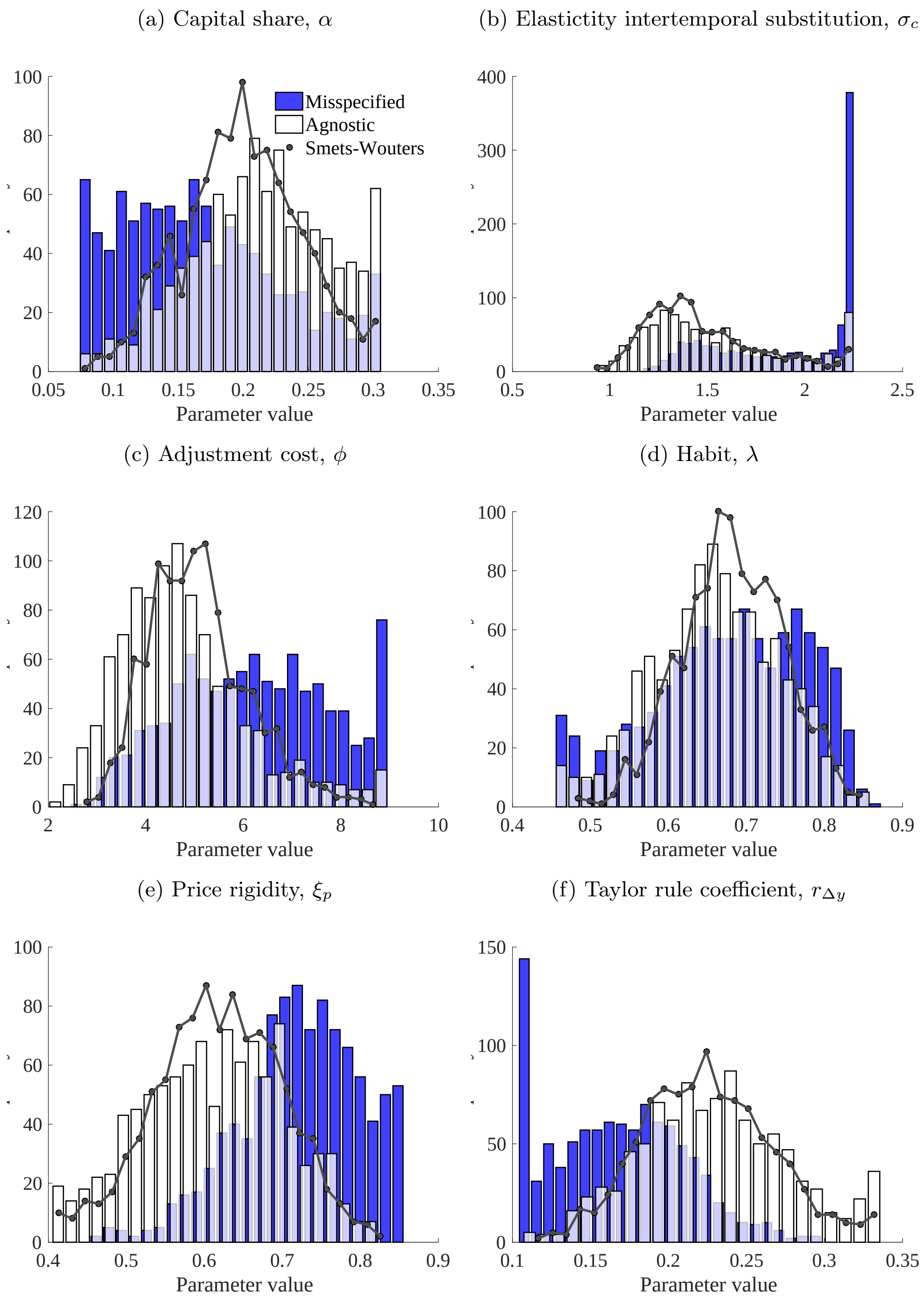
(e) Price rigidity, ξ_p

(f) Taylor rule coefficient, $r_{\Delta y}$

Note: The panels plot the distribution of the indicated parameter across the Monte Carlo replications. The color of the histograms for the misspecified case changes in a lighter shade when they overlap with the histogram for the agnostic specification. In this experiment, the true *dgp* does not include the monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

The associated mean is equal to 0.09, whereas the true value is equal to 0.19. The figure also documents that a large number of estimates are clustered at the imposed lower bound. That is, by imposing bounds we limited the distortions due to misspecification. For α, the leftward shift is so large, that there is little overlap between the distribution of the estimates based on the misspecified model and the other two empirical models. Bunching at the lower or upper bound is more pervasive for the

Figure 3.3: Histograms: Second Monte Carlo experiment



Note: The panels plot the distribution of the indicated parameter across the Monte Carlo replications. The color of the histograms for the misspecified case changes in a lighter shade when they overlap with the histogram for the agnostic specification. In this experiment, the true *dgp* does not include the TFP disturbance, but the empirical model leaves out the investment disturbance instead.

first experiment, but also observed for the second.

For the parameters considered in these figures, the distribution of estimates for the agnostic and the fully-specified SW specification are almost always centered around the true parameter value. In principle, there could be a small sample bias, since this is a complex nonlinear estimation problem. The full set of results, discussed in appendix 3.7.3, do indeed indicate that there is a bias for some parameters. In those cases, the bias is similar for the estimator based on the fully-specified specification and the agnostic one. An example of a parameter that is estimated with bias is the labor supply elasticity with respect to the real wage, σ_l. Its true value is equal to 1.92. In the first experiment, the average estimate across the Monte Carlo replications is equal to 1.84 for the SW and 1.71 for the agnostic specification. By contrast, the associated average estimate is equal to 0.27 for the misspecified model, which indicates a large bias.

3.6 Are the SW disturbances the right ones for US data?

The Monte Carlo experiment of section 3.5 documents that the ASD procedure is a powerful tool to detect and correct for misspecification when the SW model is used as the true *dgp*. In this section, we use the ASD procedure on actual data. Specifically, we first use the ASD procedure to test the restrictions imposed by structural disturbances in the SW model using the same US postwar data as in the original SW paper. We will document that the restrictions imposed by the risk premium and the investment-specific technology disturbance are rejected by the ASD procedure. That is, replacement of these regular structural disturbances by an agnostic structural disturbance leads to an increase in the marginal data density. The restrictions of the other five disturbances are not rejected. Next, we use model selection procedures to determine the number of ASDs to include and to construct a more concise specification that excludes the agnostic disturbances from *some* model equations. The best specification obtained from these selection procedures is one with three ASDs. To conclude, we interpret the nature of these three agnostic structural disturbances by examining the sign and magnitude of the associated coefficients in model equations and the IRFs of the agnostic disturbances.

3.6.1 Introducing ASDs into the Smets-Wouters model

The ASD procedure can be used in frequentist and Bayesian settings. Since SW use a Bayesian estimation procedure, we will do the same. To estimate the model with agnostic disturbances, we use the formulation of the procedure as described in section 3.3.2, which entails adding the agnostic disturbance to model equations without restricting its impact.[44] This only requires a minor modification of the Dynare program that estimates the model for the original SW specification. We do not include agnostic disturbances in equations that define observables.[45] This means

[44]The SW model has an output gap measure that depends on the outcomes of a hypothetical parallel economy with flexible prices. If an equation in the sticky-price part of the model has an associated equation in the flexible-price part of the model, then we assume that the agnostic disturbance enters the two equations with the same reduced-form coefficient.

[45]For example, the SW specification uses consumption growth as an observable and has an equation that defines consumption growth. Allowing an agnostic disturbance to affect this equation would capture measurement error (which would be correlated with structural disturbances if this ASD also appears in other model equations with a non-zero coefficient). We do not explore this possibility to keep the analysis parsimonious and to stay close the SW approach, which does not allow for measurement error.

that there are thirteen coefficients to measure the impact of an agnostic disturbance on the system. We can normalize the standard deviation of the agnostic disturbances to one, since its coefficients are of a reduced-form nature. Thus, the difference in the number of parameters between the most general agnostic specification considered and the original SW specification is equal to twelve times the number of ASDs that are introduced. Details are given in appendix 3.7.4.

The priors for the structural parameters are identical to the ones used by SW. The prior for each agnostic coefficient is a Normal with a mean equal to what the coefficient would be according to the SW restrictions and the SW prior means.[46] By centering the priors of the agnostic coefficients around the SW restrictions, we favor the SW specification. However, the means of these priors hardly matter and our results are robust to setting the prior mean equal to zero for all coefficients.[47]

The standard deviation of the prior distribution is set equal to 0.5. This implies a very uninformed prior, since the model is linear in log variables. As a robustness check we also consider a standard deviation equal to 0.1 and we find very similar results.

3.6.2 Smets-Wouters structural disturbances: Specification tests

A specification that replaces a regular structural disturbance with an agnostic one encompasses the original specification which gives it an advantage in terms of achieving a better fit. The additional parameters, however, act as a penalty term in the marginal data density. Table 3.5 reports the marginal data densities for the different specifications. The first row reports the marginal data density for the original SW specification with its seven regular structural disturbances. The seven subsequent rows give the results when the indicated regular structural disturbances is replaced by an ASD.

Overall, these results are quite supportive of the original SW specification as the SW restrictions are preferred for five of the seven structural disturbances.[48] But the results for the risk-premium and the investment specific disturbance indicate that improvement is possible.[49]

[46]For example, suppose we use the ASD procedure to test the restrictions of the risk-premium disturbance by replacing it with an ASD. The risk-premium disturbance appears in two equations, namely the consumption/bond Euler equation and the capital-valuation equation. The prior means of the reduced-form agnostic coefficients for these two equations are set equal to the values according to the SW restrictions with structural parameters evaluated at their prior means. The reduced-form coefficients associated with the other equations have a prior mean equal to zero.

[47]Having a non-zero prior has a practical advantage. The signs of the coefficients of an agnostic disturbance are not identified. That is, one can switch the signs of the coefficients of an ASD as long as one does it for all coefficients. A necessary consequence of its agnostic nature is that the sign of an ASD disturbance has no a priori meaning. If the prior means of all ASD coefficients are zero, then the ASD coefficients can flip sign for different runs of the MCMC procedure.

[48]However, it is possible that the SW specification would be rejected against more concise agnostic disturbances, that is specifications that exclude the agnostic disturbance from some equations.

[49]When we narrow the prior of the agnostic coefficients by reducing the standard deviation to 0.1, then the restrictions of the monetary policy disturbance are also rejected. But the increase in the marginal data density is relatively small, namely from -922.40 to -920.82.

Table 3.5: MISSPECIFICATION TESTS FOR ORIGINAL SW EMPIRICAL MODEL

structural SW disturbance excluded	SAD added	marginal data density
None (original SW)	no	-922.40
TFP, ε_t^a	yes	-931.21
Risk premium, ε_t^b	yes	**-908.79**
Government expenditure, ε_t^g	yes	-934.14
Investment-specific, ε_t^i	yes	**-919.81**
Monetary policy,ε_t^r	yes	-926.88
Price mark-up, ε_t^p	yes	-938.85
Wage mark-up, ε_t^w	yes	-947.31

Note: The table reports the marginal data density for different empirical specifications. The first row reports the value for the original SW specification. The specifications considered in subsequent rows replace the indicated structural disturbance with an agnostic structural disturbance. The bold numbers indicate the cases for which the MDD is higher when the indicated structural disturbance is replaced by an agnostic disturbance.

3.6.3 Which regular and agnostic disturbances to include?

The results do not necessarily imply that we should exclude the structural risk-premium and investment disturbance. After all, it is possible that a model that includes agnostic disturbances *as well as* these two SW structural disturbances has an even higher marginal data density. To investigate this issue, we compare a set of models that do or do not include the risk-premium disturbance, that do or do not include the investment disturbance, and that include one, two, or three ASDs.[50]

Table 3.6: MODEL SELECTION PROCEDURE FOR SW MODEL: STEP 1

regular structural		agnostic			marginal
ε_t^b	ε_t^i	$\widetilde{\varepsilon}_t^A$	$\widetilde{\varepsilon}_t^B$	$\widetilde{\varepsilon}_t^C$	data density
no	no	yes	yes	no	**-906.85**
no	no	yes	yes	yes	-925.55
no	yes	yes	no	no	-908.79
no	yes	yes	yes	no	-907.46
no	yes	yes	yes	yes	-922.94
yes	no	no	yes	no	-919.81
yes	no	yes	yes	no	-907.32
yes	no	yes	yes	yes	-921.71
yes	yes	no	no	no	-922.40
yes	yes	yes	no	no	-909.35
yes	yes	no	yes	no	-920.26
yes	yes	yes	yes	no	-908.09
yes	yes	yes	yes	yes	-922.82

Note: The table reports the marginal data density for different empirical specifications regarding three agnostic disturbances and the two disturbances that are misspecified, that is, the risk-premium disturbance, ε_t^b, and the investment disturbance, ε_t^i. The number in bold indicates the highest outcome.

Table 3.6 reports the results. It shows that the model with the highest marginal data density is one with two agnostic disturbances and without the SW risk-premium

[50]To estimate the model with all seven observables, an empirical specification with only one ASD would need either the risk-premium or the investment disturbance to avoid a singularity.

as well as the SW investment-specific disturbance. Another indication that there is no need for these two SW structural disturbances is that their role in terms of explaining variation in the data is very small when agnostic disturbances are included. According to the (unconditional) variance decomposition of the estimated SW model, the risk-premium disturbance is especially important for the price of capital, consumption growth, and output growth explaining 45.4%, 61.2%, and 22.1% of total variability, respectively. It only plays a minor role for other variables. When agnostic disturbances are added, then these three numbers drop to 3.88%, 3.88%, and 2.05%, respectively.[51] The reduction in the role of the investment disturbance is even stronger. In the SW model, the investment disturbance plays a quantitatively important role for many variables. For investment growth it even explains 82.1% of the volatility. With agnostic disturbances added, its role becomes minuscule. Even for investment growth it only explains 0.31%.

3.6.4 Finding the best agnostic empirical specification

To interpret ASDs, we could use the best specification found so far. However, interpretation of an ASD is easier when the specification is more concise. To determine whether an agnostic disturbance should be excluded from some equations, we implement model selection procedures using the marginal data density as the criterion of fit. This statistic increases when fit improves, but also penalizes additional parameters.

We consider both a specific-to-general procedure and a general-to-specific procedure and we apply the procedure for the specifications with two and three ASDs.[52,53] The specific-to-general procedure with three ASDs leads to the highest MDD and the selected outcome is our preferred empirical model. The specific-to-general procedure with two ASDs and the general-to-specific procedure with two ASDs lead to slightly lower MDDs.[54] Moreover, the models selected by these three procedures are very similar. Specifically, the additional ASD in the specification with three ASDs only plays a minor role. The zero restrictions imposed for the other two ASDs are not exactly the same, but the differences are due to coefficients that turn out to be small. As documented in appendix 3.7.4, the estimates of the parameters are similar and the estimates obtained with these three empirical specifications imply similar model properties. The general-to-specific procedure with three ASDs leads

[51]These numbers are based on the specification with two ASDs and all seven SW structural disturbances using posterior mode estimates.

[52]See appendix 3.7.4 for details.

[53]An informal alternative selection procedure would be the following. One starts at the same point as the general-to-specific procedure, that is, with ASDs included in every equation. The marginal posteriors of the agnostic coefficients provides information on the lack of importance of different agnostic coefficients and may provide the researcher promising combinations of zero restrictions to impose. In fact, the posteriors for the coefficients with the fully unrestricted ASD specifications are very predictive of the equations selected by the specific-to-general procedures for this application. Of course, there are good reasons why this informal procedure is not a generally accepted model selection procedure and we cannot expect this to always work well.

[54]The specific-to-general procedure generates an MDD equal to −892.92 with two ASDs and −890.76 with three. The general-to-specific with two ASDs results in an MDD of −894.94.

to a specification that has a much lower MDD.[55]

In our preferred specification, the first agnostic disturbance enters eight of the thirteen equations, the second in three, and the third in five. By contrast, the original SW risk-premium and the investment-specific disturbance appear in only two. In the remainder of this section, we discuss the estimation results for our preferred specification and give an interpretation to the three ASDs.

3.6.5 Impact on parameter estimates and model properties

As documented in appendix 3.7.4, Table 3.11, there are several differences between the estimated values of the structural parameters obtained with the fully structural SW specification and our preferred agnostic specification with three ASDs. For example, the inflation coefficient in the Taylor rule is equal to 2.05 in the SW specification and 1.77 in ours.[56] The SW estimate is right at the upper bound of our 90% highest posterior density (HPD) interval. The SW mean estimate for the parameter characterizing the share of fixed cost in production is equal to 1.61 which is quite a bit higher than our mean estimate of 1.47 and outside our 90% HPD interval. Also, the mean posterior value of the MA coefficient of the wage mark-up disturbance is equal to 0.85 according to the SW specification and 0.59 according to ours. Our mean estimate for the standard deviation of this disturbance is roughly a third of the SW estimate.

Although there are some nontrivial differences, they are relatively small and the IRFs of the five regular structural disturbances that are included in both specifications are very similar for the two empirical models. The same is true when we consider the role of these five disturbances for the variance decomposition. Details are given in appendix 3.7.4, tables 3.12 and 3.13. One nontrivial change is the role of the productivity disturbance for output growth, which is 16.1% according to SW and 22.2% according to ours. Although the differences seem minor if we consider the five structural disturbances in isolation, the combined role changes quite a bit for some variables. For example, the combined role of these five structural disturbances for investment (amount of capital used) is equal to 55.5% (74.1%) for the SW specification and 68.7% (92.6%) for our preferred specification.

3.6.6 Giving the ASDs an economic interpretation

ASDs are agnostic by nature. The model selection procedure also does not use any economic reasoning. Here we will show how the estimation results, such as parameter estimates of ASD coefficients and IRFs, can be used to give a meaningful interpretation to the ASDs. We will argue that one ASD can be interpreted

[55]Namely, -909.48. The general-to-specific procedure already stops after two steps. That is, the procedure does not detect that imposing *multiple* restrictions *simultaneously* does lead to substantial improvements. One has to impose some structure on any model selection procedure, because it would be impossible to consider all possible combinations. That is, one has to give instructions on what paths to follow and which ones to ignore. But this means that the model selection procedure may not find the best model. This motivates our use of different model selection criteria.

[56]We report posterior mean estimates unless indicated otherwise.

as an investment-specific disturbance, but with some quite striking differences from the regular one used in the literature and in SW. We will refer to this ASD as the agnostic "investment-modernization disturbance." The second ASD has features in common with the SW risk-premium disturbance, although it is closer to a preference disturbance. Moreover, like the first ASD it displays some striking differences with its original SW counterpart. We will refer to this ASD as the agnostic "Euler disturbance". The role of the third ASD is quantitatively less important than the other two. It mainly affects wage growth and is associated with a more efficient use of capital. We will refer to this ASD as the "capital-efficiency wage mark-up disturbance." By assigning names to agnostic disturbances, we may open ourselves to criticism. Our main reason for assigning these labels is that we want to make clear that agnostic disturbances are in principle theory-free, *and yet* allow the researcher to go one step further, towards giving an economic interpretation to them.

Table 3.7: ROLE OF STRUCTURAL DISTURBANCES FOR VARIANCE

	risk/preference		investment		
	SW ε_t^b	agnostic ε_t^A	SW ε_t^i	agnostic ε_t^B	agnostic ε_t^C
output	1.53	1.14	7.34	2.17	0.28
flex. price output	0	2.08	5.39	1.02	0.36
consumption	2.18	1.51	2.83	0.49	0.25
investment	0.22	1.06	44.2	29.3	1.00
hours	2.52	1.29	8.15	4.97	2.03
capital	0.04	0.12	32.5	2.37	9.75
utilization	0.86	4.14	35.4	9.46	14.7
price of capital	45.4	18.6	36.0	31.6	7.21
marginal cost	0.87	15.2	3.11	2.61	5.13
policy rate	7.40	17.2	18.3	12.5	0.65
inflation	0.58	0.68	3.18	3.96	0.91
output growth	22.1	21.3	15.8	8.04	1.82
consumption growth	61.2	61.7	0.95	2.03	0.10
investment growth	2.46	12.6	82.1	70.0	0.81

Note: The table reports the percentage of total variability explained by the SW and the agnostic risk-premium disturbance and the SW and the agnostic investment disturbance. The numbers for the SW disturbance are from estimation of the original SW model. The numbers for the agnostic disturbance are from our preferred empirical model with three agnostic disturbances.

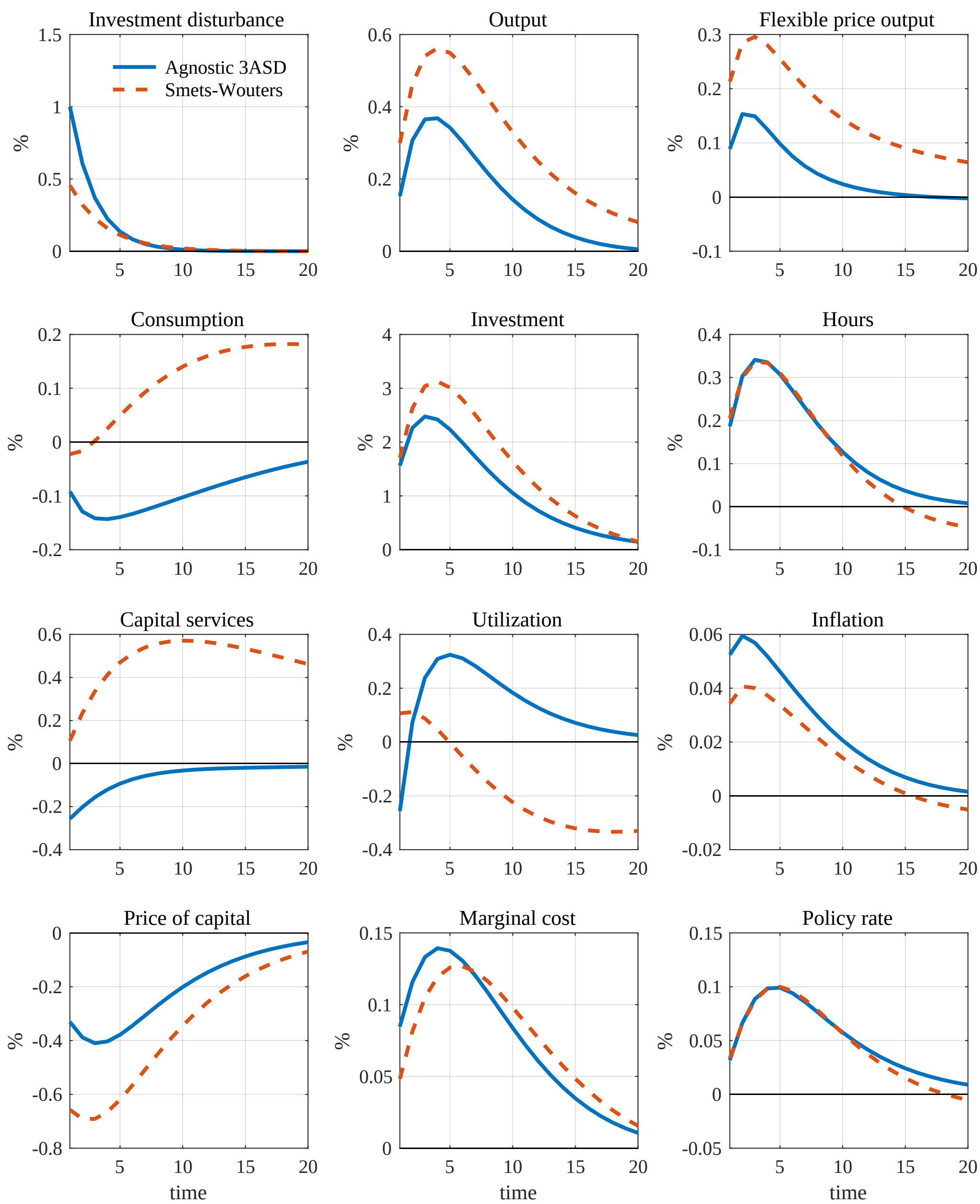
The agnostic investment-modernization disturbance, $\widetilde{\varepsilon}_t^B$

In the SW model, the investment-specific technology disturbance shows up in the investment Euler equation and in the capital accumulation equation. One of our agnostic disturbances, $\widetilde{\varepsilon}_t^B$, also shows up in these two equations.[57] The only other equation in which $\widetilde{\varepsilon}_t^B$ appears is the utilization equation that relates capacity utilization to the rental rate of capital. These findings indicate that $\widetilde{\varepsilon}_t^B$ could be interpreted as an investment-specific productivity disturbance. Furthermore, as documented in

[57]In our computer programs, the ASDs are referred to as agnA, agnB, and agnC. The economic story we are going to tell works best if we start with agnB. Labels for agnostic disturbances are arbitrary and we could relabel this disturbance as $\widetilde{\varepsilon}_t^A$, which may seem more logical given that it is discussed first. We chose not to do so, because it would create an inconsistency with our computer programs.

Table 3.7, $\tilde{\varepsilon}_t^B$, plays an important role for the volatility of investment. Specifically, it explains 70% of the volatility of investment growth compared to 82.1% for the investment-specific disturbance in the SW model. Interestingly, $\tilde{\varepsilon}_t^B$ is not important for the volatility of capital. Specifically it only explains 2.37% of the volatility of the capital stock, whereas the SW investment disturbance explains 32.5%. Thus, if $\tilde{\varepsilon}_t^B$ is an investment-specific disturbance, then it is not a typical one.

Figure 3.4: IRFs of the SW investment and corresponding agnostic disturbance



Note: These figures plot the IRFs of the SW investment-specific productivity disturbance and the agnostic disturbance $\tilde{\varepsilon}_t^B$ that we interpret as an investment-modernization disturbance.

Figure 3.4 plots the IRFs of our agnostic disturbance and the SW investment-specific disturbance. This graph documents that there are some remarkable differences. The SW investment disturbance generates a typical business cycle with key aggregates moving in the same direction. A positive agnostic investment disturbance also leads to a strong comovement between output and investment. However, a positive shock leads to a *reduction* in consumption and capital.[58] Also, whereas capacity utilization decreases in the SW model, our specification indicates an increase.

To understand these differences and to explain why we still think that $\widetilde{\varepsilon}_t^B$ is an investment-specific disturbance, we have to take a closer look at the relevant equations and how $\widetilde{\varepsilon}_t^B$ affects these equations differently than the SW investment specific disturbance, ε_t^i. The three relevant equations are the following:[59,60]

Smets-Wouters investment-specific disturbance, $\boldsymbol{\varepsilon_t^i}$

$$\begin{aligned} \text{Inv. Euler:} \quad & i_t = i_1(\Psi)\, i_{t-1} + (1 - i_1(\Psi))\, \mathbb{E}_t[i_{t+1}] + \boldsymbol{\varepsilon_t^i}, && (3.23) \\ \text{Utilization:} \quad & z_t = z_1(\Psi)\, r_t^k, && (3.24) \\ \text{Capital:} \quad & k_t = k_1(\Psi)\, k_{t-1} + (1 - k_1(\Psi))\, i_t + \boldsymbol{k_2(\Psi)\, \varepsilon_t^i}, \quad \boldsymbol{k_2(\Psi) > 0}. && (3.25) \end{aligned}$$

Agnostic investment-modernization disturbance, $\boldsymbol{\widetilde{\varepsilon}_t^B}$

$$\begin{aligned} \text{Inv. Euler:} \quad & i_t = i_1(\Psi)\, i_{t-1} + (1 - i_1(\Psi))\, \mathbb{E}_t[i_{t+1}] + \boldsymbol{d_3^B \widetilde{\varepsilon}_t^B}, \boldsymbol{d_3^B > 0}, && (3.26) \\ \text{Utilization:} \quad & z_t = z_1(\Psi)\, r_t^k + \boldsymbol{d_7^B \widetilde{\varepsilon}_t^B}, \quad \boldsymbol{d_7^B < 0}, && (3.27) \\ \text{Capital:} \quad & k_t = k_1(\Psi)\, k_{t-1} + (1 - k_1(\Psi))\, i_t + \boldsymbol{d_8^B \widetilde{\varepsilon}_t^B}, \quad \boldsymbol{d_8^B < 0}. && (3.28) \end{aligned}$$

The reason for the striking differences between the IRFs of our ASD and the SW investment disturbance is that our unrestricted approach lets the agnostic investment specific disturbance appear in the capital accumulation equation without restrictions. That is, the sign of the coefficient of $\widetilde{\varepsilon}_t^B$, d_8^B, is unrestricted, but the coefficient of ε_t^i in the SW specification, $k_2(\Psi)$ is restricted by the values of the structural parameters, Ψ. The outcome is that the posterior mean of d_8^B has the *opposite* sign relative to $k_2(\Psi)$.[61]

This means that a reduction in the cost of transforming *current* investment into

[58] Justiano et al. (2010) also report a negative consumption response to an investment disturbance, but only for the first five periods. As discussed in Ascari et al. (2016), most models would predict a countercyclical consumption response to an investment disturbance. The SW model overturns this property due to a sufficiently high degree of price and wage stickiness. Our agnostic approach implies similar estimates for price and wage stickiness, but nevertheless indicates that the data actually prefer a countercyclical consumption response.

[59] These are equations (3), (7), and (8) in the original SW paper, respectively. Ψ is the vector with structural coefficients and these restrict the coefficients in the model equations. See Smets and Wouters (2007) for the definitions of the coefficient functions. The subscripts of the coefficients of the agnostic disturbance refer to the SW equation number. For example, $d_3^B \widetilde{\varepsilon}_t^B$ is the term added to equation (3) of SW. i_t is the investment level, r_t^k the rental rate of capital, z_t the utilization rate, ε_t^i the SW investment-specific investment disturbance, $\widetilde{\varepsilon}_t^B$ the agnostic disturbance, and Ψ is the vector with structural parameters. Variables are defined relative to their steady-state values.

[60] The other two ASDs also enter some of these equations. We leave these terms out for transparency reasons and because there are no interactions in a linear framework.

[61] Moreover, the 90% HPD does not include 0. Although it does not make much of a difference, we give the SW outcome the best possible chance by setting the prior means of the coefficients of $\widetilde{\varepsilon}_t^B$ to what they would be under the SW specification using SW prior means.

capital goes together with increased depreciation of the *existing* capital stock in our specification. In the SW model, an investment specific disturbance does not affect the economic viability of the existing capital stock. Our agnostic approach questions this assumption and suggests that the investment-specific productivity disturbance goes together with scrapping of older vintages. This is the reason why we refer to it as an agnostic investment-modernization disturbance.

In the SW model, capacity utilization is proportional to the rental rate and there are no shocks that can affect this relationship. An accelerated depreciation of the capital stock increases the rental rate, which in turn would induce an increase in the utilization rate. In our agnostic specification, this relationship is dampened somewhat, since a positive agnostic disturbance has a *direct* negative impact on capacity utilization since it enters the capacity utilization with a negative coefficient. The overall effect is still an increase in capacity utilization. It seems plausible that scraping of old vintages goes together with higher utilization of the remaining capital stock.

The agnostic Euler disturbance, $\widetilde{\varepsilon}_t^A$

The agnostic disturbance $\widetilde{\varepsilon}_t^A$ appears in eight equations. This leaves open many possible interpretations. The key equation, however, is the Euler equation for bonds, because excluding the disturbance from this equation leads to by far the largest drop in the marginal data density. This suggests that it could have key characteristics in common with a preference or a risk-premium disturbance. This view is also supported by Table 3.7 which documents that $\widetilde{\varepsilon}_t^A$ is important for the same variables as the SW risk-premium disturbance. However, this agnostic disturbance also has some quite different characteristics from both. Therefore, we will adopt an alternative name and refer to it as the agnostic Euler disturbance. For the interpretation of $\widetilde{\varepsilon}_t^A$, it is important to understand the differences in impact of a regular preference and a regular (bond) risk-premium disturbance.

Difference between a preference and (bond) risk-premium disturbance. A preference disturbance affects current utility. This means it affects the marginal rate of substitution and, thus, *all* Euler equations. Such a preference disturbance is used in Smets and Wouters (2003). By contrast, Smets and Wouters (2007) include instead a (bond) risk premium that introduces a wedge between the policy rate and the required rate of return on bonds without affecting other Euler equations. Both disturbances have a strong impact on current consumption. However, a positive preference disturbance makes current consumption more desirable and reduces the attractiveness of *all* types of saving. A positive risk-premium disturbance only makes savings in bonds less attractive. That is, it induces a desire to substitute out of bonds and into investment, in addition to an increase in consumption. Thus, a preference disturbance leads to a negative comovement of consumption and investment, whereas a (bond) risk-premium disturbance leads to a positive comovement. Smets and

Wouters (2007) mention this as the reason for using a risk-premium instead of a preference disturbance.

There is another key difference between these two disturbance. A preference disturbance affects output in both the flexible-price and the sticky-price part of the model. By contrast, a risk-premium disturbance has no affect on key aggregates such as consumption and output in the flexible price part of the SW model.[62]

Is $\widetilde{\varepsilon}_t^A$ a preference, a risk-premium, or another type of disturbance? Figure 3.5 plots the IRFs of the SW risk-premium and our agnostic disturbance. The figure documents that both generate a regular business cycle with positive comovement for output, consumption, investment, and hours. The positive comovement suggest that the agnostic disturbance is a bond risk-premium disturbance as in Smets and Wouters (2007) and not a preference disturbance as in Smets and Wouters (2003). However, the agnostic disturbance has a strong impact on flexible-price output which is inconsistent with it being a (bond) risk-premium disturbance and consistent with it being a preference disturbance. These differences are big enough for us to come up with a new label and we choose Euler disturbance.

To better understand the nature of the agnostic Euler disturbance, we take a closer look at the equations in which $\widetilde{\varepsilon}_t^A$ enters. It appears in the aggregate budget constraint, the bond Euler equation, the investment Euler equation, the capital value equation, the utilization rate equation, the price mark-up equation, the rental rate of capital equation, and the Taylor rule.

Although our agnostic disturbance does have some effect on quite a few different aspects of the model, the interpretation is eased by the fact that the role of the agnostic disturbance is minor in most of the eight equations in the sense that allowing it to enter these equations only has a minor quantitative impact on the behavior of model variables or only affects the qualitative behavior of one or two variables without affecting the behavior of the key macroeconomic variables.

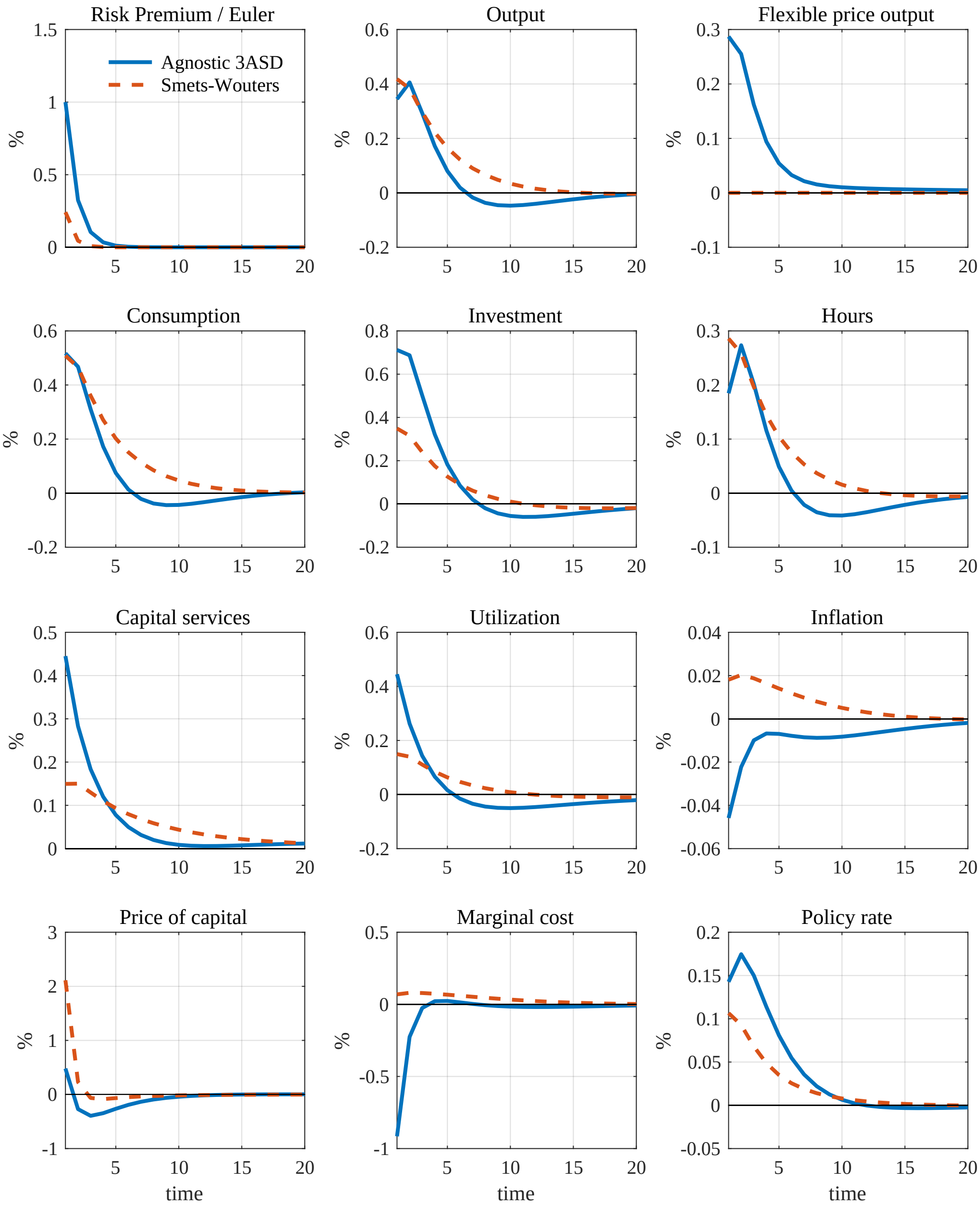
Specifically, to understand the role of $\widetilde{\varepsilon}_t^A$ on key macroeconomic aggregates we can restrict ourselves to the Taylor rule and the three model equations that are relevant for the savings/investment decisions, which are the bond Euler equation, the investment Euler equation, and the capital value equation. The following set of equations documents how the SW risk-premium disturbance and our agnostic Euler enter these equations:[63]

---

[62]The reason is the following. In the flexible price part of the model, the nominal policy rate, r_t, the expected inflation rate, $\mathbb{E}_t[\pi_{t+1}]$, and the risk-premium disturbance, ε_t^b, only appear in the combination $r_t - \mathbb{E}_t[\pi_{t+1}] + \varepsilon_t^b$. Consequently, a change in ε_t^b is simply absorbed by the real rate. This is not the case in the sticky-price economy, because it would be inconsistent with the Taylor rule.

[63]In these equations, c_t is consumption, l_t is hours worked, r_t is the nominal policy rate, π_t is the inflation rate, q_t is the price of capital, y_t is output, and y_t^p is output in the flexible-price economy. Also see information given in footnote 59.

Figure 3.5: IRFs of the SW risk-premium and the agnostic Euler disturbance



Note: These figures plot the IRFs of the SW risk-premium disturbance and the agnostic disturbance $\tilde{\varepsilon}_t^A$ that we interpret as an Euler-equation disturbance.

Smets-Wouters risk premium, ε_t^b

$$\begin{aligned}
\text{Bond Euler:} \quad & c_t = c_1(\Psi)\, c_{t-1} + (1 - c_1(\Psi))\, \mathbb{E}_t[c_{t+1}] \\
& + c_2(\Psi)\,(l_t - \mathbb{E}_t[l_{t+1}]) \\
& - c_3(\Psi)\, \boldsymbol{(r_t - \mathbb{E}_t[\pi_{t+1}] + \varepsilon_t^b)}, c_3(\Psi) > 0, && (3.29) \\
\text{Inv. Euler:} \quad & i_t = i_1(\Psi)\, i_{t-1} \\
& + (1 - i_1(\Psi))\, \mathbb{E}_t[i_{t+1}] + \varepsilon_t^i, && (3.30) \\
\text{Capital value:} \quad & q_t = q_1 \mathbb{E}_t[q_{t+1}] + (1 - q_1)\mathbb{E}_t\left[r_{t+1}^k\right] \\
& \boldsymbol{-(r_t - \mathbb{E}_t[\pi_{t+1}] + \varepsilon_t^b)}, && (3.31) \\
\text{Policy rate:} \quad & r_t = \rho r_{t-1} + (1 - \rho)\{r_\pi + r_Y(y_t - y_t^p)\} \\
& + r_{\Delta y}[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \varepsilon_t^r. && (3.32)
\end{aligned}$$

Agnostic Euler disturbance, $\widetilde{\varepsilon}_t^A$

$$\begin{aligned}
\text{Bond Euler:} \quad & c_t = c_1(\Psi)\, c_{t-1} + (1 - c_1(\Psi))\, \mathbb{E}_t[c_{t+1}] \\
& + c_2(\Psi)\,(l_t - \mathbb{E}_t[l_{t+1}]) \\
& - c_3(\Psi)\, \boldsymbol{(r_t - \mathbb{E}_t[\pi_{t+1}]) - d_2^A \widetilde{\varepsilon}_t^A, d_2^A > 0}, && (3.33) \\
\text{Inv. Euler:} \quad & i_t = i_1(\Psi)\, i_{t-1} \\
& + (1 - i_1(\Psi))\, \mathbb{E}_t[i_{t+1}] + \varepsilon_t^i \boldsymbol{- d_3^A \widetilde{\varepsilon}_t^A, d_3^A > 0}, && (3.34) \\
\text{Capital value:} \quad & q_t = q_1 \mathbb{E}_t[q_{t+1}] + (1 - q_1)\mathbb{E}_t\left[r_{t+1}^k\right] \\
& \boldsymbol{-(r_t - \mathbb{E}_t[\pi_{t+1}]) - d_4^A \widetilde{\varepsilon}_t^A, d_4^A > 0}, && (3.35) \\
\text{Policy rate:} \quad & r_t = \rho r_{t-1} + (1 - \rho)\{r_\pi + r_Y(y_t - y_t^p)\} \\
& + r_{\Delta y}[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \varepsilon_t^r \boldsymbol{+ d_{14}^A \widetilde{\varepsilon}_t^A, d_{14}^A > 0}. && (3.36)
\end{aligned}$$

As in SW, we use the bond Euler equation to substitute the marginal rate of substitution out of the capital valuation equation. While the SW bond risk-premium disturbance, ε_t^b, does *not* appear in the original capital valuation equation, it does show up *after* this substitution has taken place. Moreover, it appears in these two equations with the exact same coefficient as the nominal interest rate for bonds, r_t. By contrast, *after* substituting out the marginal rate of substitution in the capital value equation, a preference disturbance would *no longer* appear in the capital valuation equation.[64]

Our ASD appears in the bond Euler equation and the capital valuation equation and it shows up with the same sign as the SW risk-premium disturbance. This supports the view that our ASD is similar to a risk-premium disturbance. Nevertheless,

[64]In the SW specification, the impact of the risk-premium disturbance is normalized to be equal to 1 in one of the equations. The actual impact of this disturbance on this equation is then determined by the estimated standard deviation. Instead, we normalize the standard deviation of the ASDs. We do not want to impose the SW normalization, since it would imply that the agnostic disturbance must affect the equation in which the coefficient is normalized unless the estimated standard deviation is equal to zero, which would mean that it would not have an effect on any other equation either.

one could argue that the ASD is a preference and not a bond risk-premium disturbance for the following reasons. Although d_4^A has the right sign for a risk-premium coefficient, its magnitude, evaluated using the posterior mean, is too small.[65] The 90% HPD interval of the coefficient of $\widetilde{\varepsilon}_t^A$ in the capital valuation equation, d_4^A, includes zero and setting the coefficient equal to zero has very little impact on model properties and virtually none on the marginal data density.

As pointed out in Smets and Wouters (2003), a preference disturbance generates consumption and investment responses that move in opposite directions. Our ASD predicts responses in the same direction even if we impose that the ASD does not enter the capital valuation equation (after substituting out the MRS). The reason for the positive comovement is that our ASD also enters the investment Euler equation. The investment Euler equation is a dynamic equation, but its dynamic aspects are due solely to investment adjustment costs.[66] Our agnostic approach indicates that the structural disturbance that plays a key role in the bond Euler equation should also appear in the investment Euler equation. In fact, it is the first equation chosen in our specific-to-general model selection procedure.

So what could this agnostic disturbance represent? The simplest – and our preferred explanation – is that it is a preference disturbance that is correlated with an investment-specific disturbance.[67] A more structural interpretation would be the following. A preference disturbance would also affect the (linearized) investment Euler equation if investment does not only lead to expenses in the current, but also in subsequent periods. For example, investment may lead to additional expenses when capital becomes productive. A positive preference disturbance would lower the value of such future liabilities.

This disturbance appears directly in the Taylor rule with a negative coefficient. This means that the central bank responds more aggressively to business cycle fluctuations induced by this Euler disturbance. Without this effect on the Taylor rule this disturbance would have a stronger impact on economic aggregates and inflation would no longer be procyclical.[68]

---

[65]If our ASD is a risk-premium disturbance, then d_4^A/d_2^A should be equal to $1/c_3(\Psi)$, but using posterior means, we find that $d_4^A/d_2^A = 3.3$, whereas $1/c_3(\Psi) = 7.27$, substantially higher. Here, c_3 is a function of the habit, the elasticity of inter-temporal substitution, and the trend growth rate parameter. c_3 is calculated using posterior means of our preferred specification.

[66]Adjustment costs are zero in the steady state, which implies that neither a preference disturbance nor a risk-premium disturbance appear in a *linearized* investment Euler equation. A preference disturbance would appear in the original *nonlinear* equation. The main intertemporal aspect of the investment decision, which is also present without adjustment costs, is captured by the capital valuation equation.

[67]As discussed above, our agnostic structural investment disturbance, $\widetilde{\varepsilon}_t^B$, enters the capital accumulation equation with a sign that is the opposite of the regular investment disturbance, which we interpreted as scrapping of older vintages. $\widetilde{\varepsilon}_t^A$ does *not* enter the capital accumulation equation. This would indicate that this investment disturbance which goes together with an upswing in agents' mood is between a regular investment disturbance and our agnostic investment disturbance in terms of what it implies for the viability or depreciation of the existing capital stock.

[68]See appendix 3.7.4.

The agnostic capital-efficiency wage mark-up disturbance, $\widetilde{\varepsilon}_t^C$

The third agnostic disturbance chosen by our model selection criterion increases the total number of structural disturbances to eight, that is, one more than the number in the SW specification. Thus, this agnostic disturbance cannot be interpreted as a replacement of a SW disturbance. Figure 3.6 plots its IRFs.

This third agnostic disturbance, $\widetilde{\varepsilon}_t^C$, appears in five equations. The first equation into which it is selected is the wage-adjustment equation. It also shows up into three equations related to capital, namely the capital accumulation equation, the capital utilization equation, and the capital-valuation equation. Finally, it appears in the economy-wide budget constraint, although the impact on the latter is minor.

The SW wage mark-up disturbance, ε_t^w also shows up in the wage-adjustment equation. The differences with $\widetilde{\varepsilon}_t^C$ are the following. First, ε_t^w *only* shows up in the wage-adjustment equation, whereas $\widetilde{\varepsilon}_t^C$ has a direct impact on key equations related to capital. This is an important difference that results in quite different IRFs. A positive shock to ε_t^w induces a regular economic downturn with all key macroeconomic aggregates moving in the same direction, except for the price of capital which increases initially. A positive shock to $\widetilde{\varepsilon}_t^C$ also induces a recession with a reduction in output, investment, and employment. However, it leads to an increase in potential output, installed capital, and initially also an increase in capacity utilization. In contrast to the SW ε_t^w shock it leads to a decrease in the price of capital.

The second difference between our agnostic $\widetilde{\varepsilon}_t^C$ and the SW ε_t^w disturbance is that a shock to $\widetilde{\varepsilon}_t^C$ is very temporary. $\widetilde{\varepsilon}_t^C$ is an AR(1) process, and the posterior mean of the auto-regressive coefficient is equal to 0.19. The SW ε_t^w disturbance is a very persistent ARMA(1,1) process. The presence of $\widetilde{\varepsilon}_t^C$ in the empirical model strongly reduces the coefficient of the MA component of ε_t^w, but has little impact on the AR component.[69]

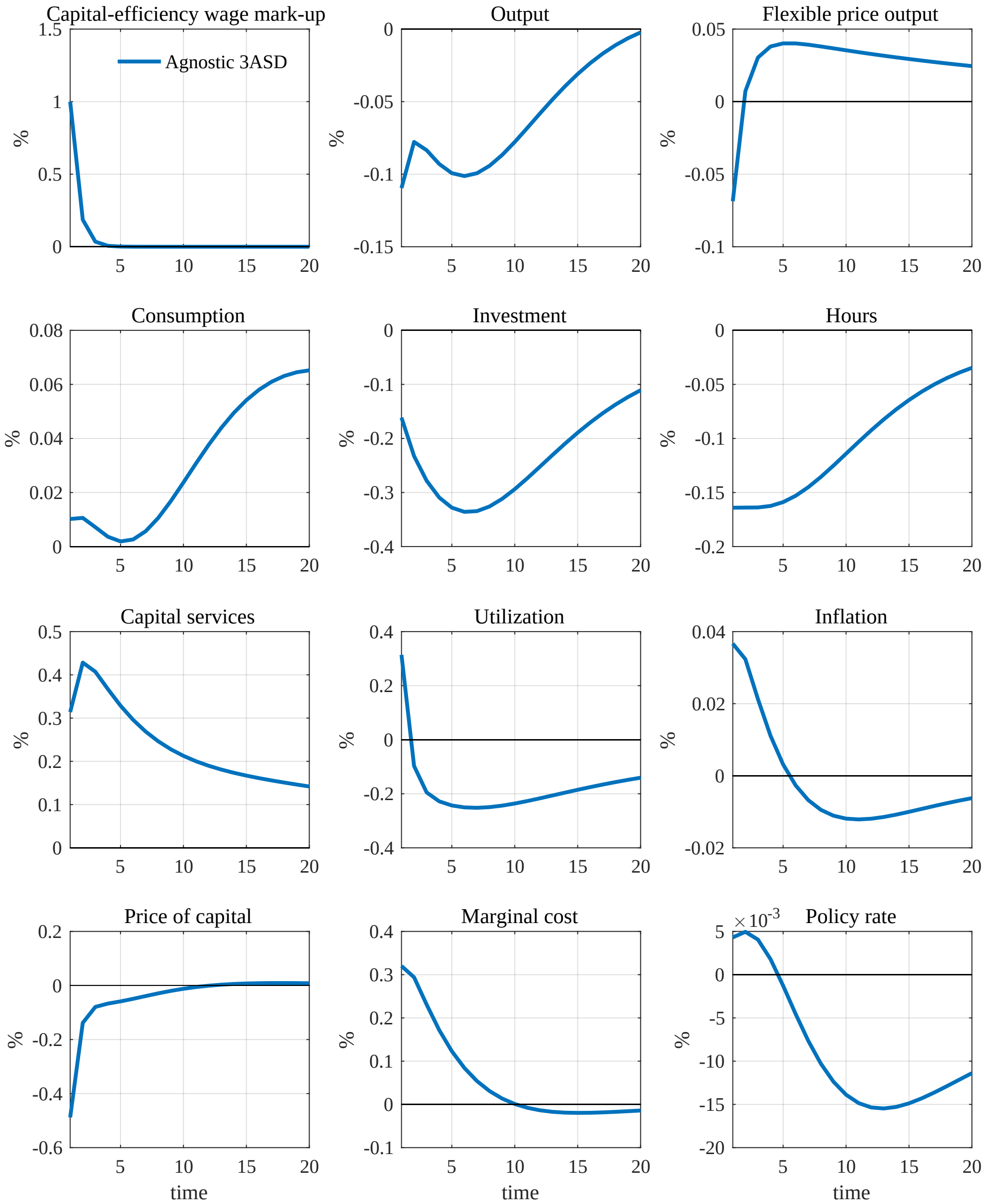
Including $\widetilde{\varepsilon}_t^C$ in the empirical specification does not reduce the role of ε_t^w for fluctuations of key variables. ε_t^w remains the most important disturbance for key economic aggregates. The only exception is the wage *growth* rate. In the SW specification ε_t^w explains 61.6% of the volatility of wage growth, whereas it only explains 13.3% in our preferred specification. This role is clearly taken over by $\widetilde{\varepsilon}_t^C$ which explains 53.5% of wage growth volatility. $\widetilde{\varepsilon}_t^C$ also plays a nontrivial role for fluctuations in the capital stock, capacity utilization, and the rental rate of capital, explaining 9.8%, 14.7%, and 13.1%, of total variability respectively.

The following equations document how $\widetilde{\varepsilon}_t^C$ enters the model:[70]

[69] Specifically, with $\widetilde{\varepsilon}_t^C$ included in the empirical specification the posterior means of the AR and the MA coefficients of ε_t^w are equal to 0.97 and 0.59, respectively. Estimates with the SW specification for these two numbers are 0.97 and 0.85.

[70] We leave out the overall budget constraint since the role of the disturbance in this equation is minor, but its impact in this equation is like a contractionary fiscal expenditure shock. Details are given in appendix 3.7.4. w_t is the real wage rate and μ_t^w is the real wage mark-up, i.e., the difference between the wage rate and the marginal rate of substitution between consumption and leisure. Also see footnote 59 for additional information.

Figure 3.6: IRFs of the agnostic capital-efficiency wage mark-up disturbance



Note: These figures plot the IRFs of the agnostic disturbance $\widehat{\varepsilon}_t^C$ that we interpret as a capital-efficiency wage mark-up disturbance.

Agnostic capital-efficiency wage mark-up disturbance, $\widetilde{\varepsilon}_t^C$

$$
\begin{aligned}
\text{Capital value: } q_t &= q_1\mathbb{E}_t\left[q_{t+1}\right] + (1-q_1)\mathbb{E}_t\left[r_{t+1}^k\right] \\
&- \left(r_t - \mathbb{E}_t\left[\pi_{t+1}\right]\right) - \boldsymbol{d_4^C \widetilde{\varepsilon}_t^C, d_4^C < 0}, && (3.37) \\
\text{Utilization: } z_t &= z_1\left(\Psi\right) r_t^k + \boldsymbol{d_7^C \widetilde{\varepsilon}_t^C, \quad d_7^C > 0}, && (3.38) \\
\text{Capital: } k_t &= k_1\left(\Psi\right) k_{t-1} + \left(1-k_1\left(\Psi\right)\right) i_t + \boldsymbol{d_8^C \widetilde{\varepsilon}_t^C, \quad d_8^C > 0}, && (3.39) \\
\text{Wage mark-up: } w_t &= w_1 w_{t-1} + (1-w_1)(\mathbb{E}_t[w_{t+1} + \pi_{t+1}] && (3.40) \\
&- w_2\pi_t + w_3\pi_{t-1} - w_4\mu_t^w + \boldsymbol{d_{13}^C \widetilde{\varepsilon}_t^C, \quad d_{13}^C > 0}. && (3.41)
\end{aligned}
$$

The equations indicate that this agnostic disturbance increases the wage mark-up and is associated with increased efficiency of the capital stock, both in terms of a lower depreciation rate and increased utilization. It also goes together with a reduction in the value of existing capital. Thus, this ASD could capture an increase in the wage rate, for example, because of increased bargaining power of workers, in response to which firms use capital more efficiently. An alternative is that its origin lies in changes in the ability or need to use capital more efficiently, but that a more efficient use of capital comes at the cost of higher wage rates. That is, to adopt this more efficient use of capital, firms have to pay a higher wage rate, perhaps in terms of an overtime premium.

3.7 Appendices

3.7.1 Consequences of misspecification: An analytical example

In this section, we give a *very* simple example to indicate that misspecification can have large distortive effects in the sense that *implied* properties of the model using the parameter estimates can be at odds with the *actual* corresponding properties of the data that are used to estimate the parameters. The model is linear, and all variables have a Normal distribution. Throughout this section, parameter estimates are based on population moments. Thus, the results are not due to small sample variation. The estimation procedure is Maximum Likelihood (ML).

More specifically, this example demonstrates that there can be massive differences between the variances of observables as *implied* by the model using estimated parameter values and the actual variances in the data set. This result is surprising since the ML estimator of the variance of a given time series is the sample variance when the variable has a Normal distribution. We will show that this is not necessarily true for implied variances when the empirical model is misspecified.[71]

True model. The true model is given by the following set of equations:

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} = \Lambda \varepsilon_t, \tag{3.42}$$

$$\mathbb{E}\left[\varepsilon_t \varepsilon_t'\right] = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}, \tag{3.43}$$

and we make the following assumption about the distribution of the error terms:

$$\varepsilon_{1,t} \sim N(0, \sigma_1^2) \text{ and } \varepsilon_{2,t} \sim N(0, \sigma_2^2). \tag{3.44}$$

Misspecification. The objective is to estimate the standard deviations of the structural disturbances, σ_1^2 and σ_2^2. The researcher takes the value of Λ as given. The empirical model is misspecified, because $\overline{\Lambda} \neq \Lambda$ is used instead of the true value.

[71] As a byproduct of this paper, we learned that there also can be large gaps between *actual* properties of the data used and the corresponding *implied* properties according to the Maximum Likelihood estimates of the model parameters when the DSGE model is correctly specified, but a data sample with finite length is used. Since the objective of Maximum Likelihood is not to match moments, there is no reason why there should be a close match, but we were surprised by the large magnitudes of the differences. For example, using a sample of 1,000 observations generated by the SW model with seven disturbances and the correct empirical specification, it is not unusual to find implied standard deviations for the observables that are three to five times their data counterpart. Such differences will disappear as the sample size increases, since the estimator is consistent, but such asymptotic results do not provide much assurance if there is a small sample bias even at a relatively large sample size of 1,000 observations.

Empirical specifications. We consider the following two empirical specifications:

Case 1: Empirical model given by

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \overline{\Lambda}\varepsilon_t, \quad \mathbb{E}\left[\varepsilon_t \varepsilon_t'\right] = \begin{bmatrix} \overline{\sigma}_1^2 & \overline{\sigma}_{12} \\ \overline{\sigma}_{12} & \overline{\sigma}_2^2 \end{bmatrix}. \tag{3.45}$$

Case 2: Empirical model given by

$$y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \overline{\Lambda}\varepsilon_t, \quad \mathbb{E}\left[\varepsilon_t \varepsilon_t'\right] = \begin{bmatrix} \overline{\sigma}_1^2 & 0 \\ 0 & \overline{\sigma}_2^2 \end{bmatrix}. \tag{3.46}$$

Both empirical models are misspecified, because they use the wrong value of Λ. In the first case, the empirical model allows the correlation between the two innovations to be non-zero even though it is equal to zero according to the true data generating process. In the second case, the empirical model imposes that the correlation is equal to zero, just as it is in the true model.

Case 1: Wrong Λ and allow for wrong σ_{12}. Since the model is linear and the shocks have a normal distribution, the ML estimator of the variance-covariance matrix $\mathbb{E}[\varepsilon_t \varepsilon_t']$, $\widehat{\Sigma}_\varepsilon$, is given by

$$\widehat{\Sigma}_\varepsilon = \overline{\Lambda}^{-1}\widehat{\Sigma}_y'\overline{\Lambda}^{-1\prime}. \tag{3.47}$$

As mentioned above, we abstract from sampling variation and $\widehat{\Sigma}_y'$ is estimated using population moments. This means that the ML estimator of $\widehat{\Sigma}_\varepsilon'$ is given by

$$\begin{aligned} \widehat{\Sigma}_\varepsilon &= \overline{\Lambda}^{-1}\mathbb{E}\left[y_t y_t'\right]\overline{\Lambda}^{-1\prime} & (3.48) \\ &= \overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime}. & (3.49) \end{aligned}$$

True versus implied variance. The purpose of this section is to document the consequences of misspecification for the implied variance of the observable y_t according to the estimated model. The *true* variance-covariance matrix is given by:

$$\Sigma_y^{\text{true}} = \mathbb{E}\left[y_t y_t'\right] = \Lambda\Lambda'. \tag{3.50}$$

The *implied* variance of y_t according the researcher's (misspecified) model, $\widehat{\Sigma}_y$, is given by

$$\begin{aligned} \widehat{\Sigma}_y &= \overline{\Lambda}\widehat{\Sigma}_\varepsilon\overline{\Lambda}' & (3.51) \\ &= \overline{\Lambda}\overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime}\overline{\Lambda}' & (3.52) \\ &= \Lambda\Lambda' = \Sigma_y^{\text{true}}. & (3.53) \end{aligned}$$

Thus, the procedure actually generates the correct answer even though an incorrect empirical specification is used. In this case, the estimated empirical model is misspecified for two reasons, namely it has the wrong Λ and the estimated value of σ_{12} is not equal to its true value. These have exactly offsetting effects in terms of their impact on the implied variance. Another way to look at this result is the following. By allowing for a more flexible specification, i.e., a non-zero value for σ_{12}, the researcher would get a better answer for the implied variance of y_t even though the flexibility implies that the estimated model is wrong in more dimensions.

Case 2: Wrong Λ and correct σ_{12}. Obtaining the estimate for $\widehat{\Sigma}_\varepsilon$ is just as easy as in the previous case. Given $\widehat{\Lambda}$ and data for y_t, one can calculate the values for ε_t and use these to calculate the variance of ε_t and the implied variance of y_t. The following is a complicated, but useful way to express the outcome:

$$\widehat{\Sigma}_\varepsilon = \begin{array}{c} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ + \\ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{array} \tag{3.54}$$

True versus implied variance. The implied variance of y_t is equal to

$$\widehat{\Sigma}_y = \left(\begin{array}{c} \overline{\Lambda} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}' \\ + \\ \overline{\Lambda} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}^{-1}\Lambda\Lambda'\overline{\Lambda}^{-1\prime} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \overline{\Lambda}' \end{array} \right) \neq \Lambda\Lambda' = \Sigma_y^{\text{true}} \tag{3.55}$$

The reason $\widehat{\Sigma}_y \neq \Sigma_y^{\text{true}}$ is that the $\overline{\Lambda}$ terms do not cancel out. In our Monte Carlo experiments with misspecified models, we find that there often are large gaps between the variances of the observables used in the estimation and the corresponding variances as implied by the model using the estimated parameters. Moreover, there is a bias. That is, the implied variance is typically larger than the actual variance. Our Monte Carlo experiments are a lot more complicated than this example, but this example may shed light on the coincidence of high implied variances. Specifically, because the $\overline{\Lambda}$s do not cancel out, the expression for $\widehat{\Sigma}_y$ contains terms like the following:

$$\overline{\Lambda} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \overline{\Lambda}^{-1} = \frac{1}{\overline{\lambda}_{11}\overline{\lambda}_{22} - \overline{\lambda}_{12}\overline{\lambda}_{21}} \begin{bmatrix} \overline{\lambda}_{11}\overline{\lambda}_{22} & -\overline{\lambda}_{11}\overline{\lambda}_{12} \\ \overline{\lambda}_{21}\overline{\lambda}_{22} & -\overline{\lambda}_{12}\overline{\lambda}_{21} \end{bmatrix}. \tag{3.56}$$

This equation documents that the ratio of the implied variance relative to the true variance could be arbitrarily large if the term in the denominator goes to zero.[72]

[72]The opposite is less likely, since it would require values for the λ_{ij} coefficients such that the combinations appearing in square brackets are small, but the particular combination in the denominator is not. For example, one cannot accomplish this by simply choosing small values for the λ_{ij} terms.

For a correctly specified model this would not matter, since the small term in the denominator would then be offset by an equally small term in the numerator. But this is not necessarily the case for an incorrectly specified model.

3.7.2 Identification of structural parameters

We use the test proposed in Komunjer and Ng (2011) to check whether the parameters of the empirical specifications used in our experiments are identified. We will refer to this test as the KN test. This test provides both necessary and sufficient conditions for local identification under a set of weak conditions.[73] It focuses on the state-space representation of the model and – in contrast to earlier identification tests – does not require the user to specify a set of particular autocovariances.[74]

Identification of original Smets-Wouters estimation exercise. SW fix the values of five parameters: depreciation, δ, steady-state wage mark-up, $\overline{\mu}$, steady-state exogenous spending, $\overline{g}$, curvature in the Kimball goods-market aggregator, ε_p, and curvature in the Kimball labor-market aggregator, ε_w. Komunjer and Ng (2011) consider the identification of the SW model, but their empirical specification is slightly different from the one of SW in that all variables are demeaned. By contrast, the data in the SW estimation exercise does contain information about the level, since the inflation rate and the nominal interest rate are in levels. Komunjer and Ng (2011) show that several subsets of the five parameter restrictions mentioned above are sufficient to obtain identification *if* the parameter controlling steady state hours, $\bar{l}$, and the parameter controlling steady state inflation, $\overline{\pi}$, are fixed as well. It makes sense that identification requires more restrictions when information about the levels is not used in the estimation.

Identification of our specifications. The empirical and true specifications used in our Monte Carlo experiments have six structural disturbances, whereas the original SW empirical model has seven. This may imply that less parameters are identified. It is important that the parameters that we try to estimate are identified. If parameters are not identified, then different parameter combinations lead to the same criterion of fit used in the estimation, so it would not be surprising if parameter estimates are different for slightly different specifications.

Consequently, we adopt the following conservative strategy to ensure identification. The KN test checks rank conditions of matrices and to see whether there is a singularity one needs to choose a tolerance criterion. We set the criterion at a level that is more strict than the one chosen by KN.[75] We follow SW and fix the values of the five parameters mentioned above. In addition, we fix all parameters that have a direct effect on the means of variables, since we use demeaned variables in the estimation. The associated parameters are the trend growth rate, $\overline{\gamma}$, the parameter

[73] These are a stability condition and regularity conditions on the innovations.
[74] An example of such an earlier test is Iskrev (2010).
[75] We set"Tol" equal to 1e-2 instead of 1e-3 (a higher number is more strict).

Table 3.8: Komunjer and NG identification test

required number		41	225	36	302	-
	n	$\overline{\Delta}_{\Lambda}^{S}$	$\overline{\Delta}_{T}^{S}$	$\overline{\Delta}_{U}^{S}$	$\overline{\Delta}^{S}$	pass?
ε_t^a excluded	12	41	225	36	302	yes
ε_t^b excluded	12	41	225	36	302	yes
ε_t^g excluded	12	41	225	36	302	yes
ε_t^i excluded	12	41	225	36	302	yes
ε_t^r excluded	12	41	225	36	302	yes
ε_t^w excluded	13	41	225	36	302	yes
ε_t^p excluded	13	41	225	36	302	yes

Note: Here, n is the number of restrictions, which includes the number of coefficients fixed in all experiments and the number of coefficients in the law of motion of the excluded exogenous random variable that are all set to zero. $\overline{\Delta}_{\Lambda}^{S}$ is a matrix that contains the derivatives of all the vectorized elements in the state-space representation of the model (the A, B, C, D matrices and the variance-covariance matrices) evaluated at the true parameter values. It is intuitive that this matrix needs to have full rank for identification. But it is not sufficient. $\overline{\Delta}_{T}^{S}$ and $\overline{\Delta}_{U}^{S}$ are matrices with particular elements related to the state-space representation. The matrix $\overline{\Delta}^{S} = [\overline{\Delta}_{\Lambda}^{S}\ \overline{\Delta}_{T}^{S}\ \overline{\Delta}_{U}^{S}]$ needs to have full rank to pass the KN test.

controlling steady state hours, $\bar{l}$, the parameter controlling steady state inflation, $\overline{\pi}$, and the discount factor, β.[76] Finally, as discussed in section 3.2.1, we fix the spillover from the productivity disturbance to exogenous spending and set it equal to zero.

The results of the KN test are reported in Table 3.8 and it indicates that the identification test is passed in all cases. That is, lack of identification is not driving the results in section 3.2.

3.7.3 Additional results for Monte Carlo experiments

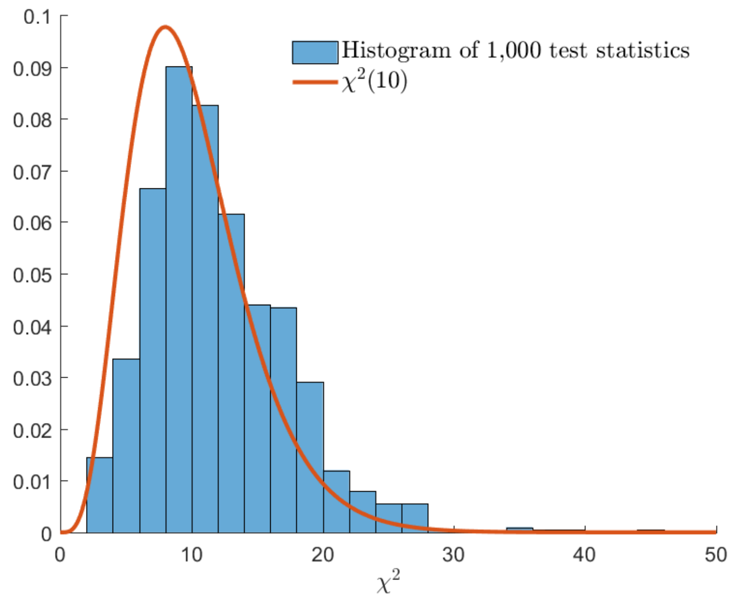
In this appendix, we report additional results for the analysis of section 3.5 in which we compared estimation outcomes using the ASD specification, the SW model with only regular structural disturbances, and an incorrect empirical model.

Figures 3.7 and 3.8 plot the histograms of the estimated χ^2 statistics across Monte Carlo replications for the two experiments of section 3.5 together with the theoretical (large-sample) χ^2 distribution. The number of degrees of freedom is equal to 10.

Tables 3.9 and 3.10 document detailed information on the distribution of parameter estimates for the two Monte Carlo experiments.

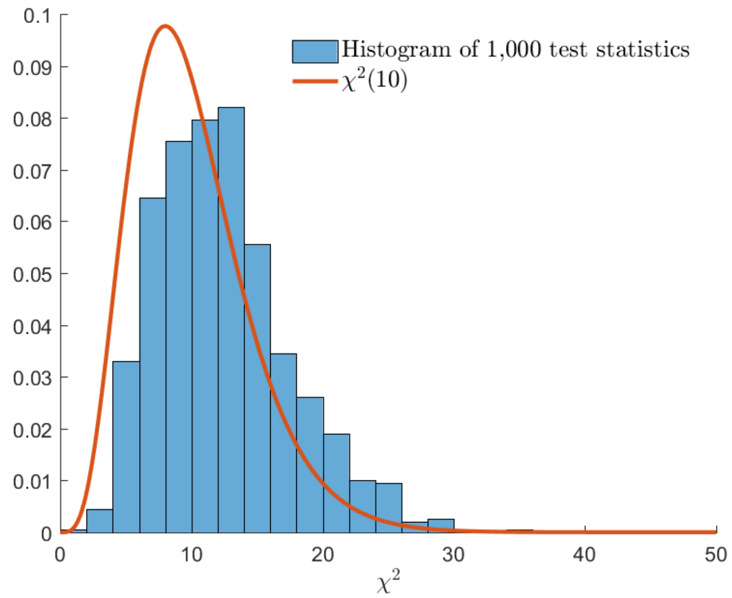
[76]It is a conservative choice to fix all four, since identification only requires that two parameters are fixed according to the test of Komunjer and Ng (2011).

Figure 3.7: Likelihood ratio test agnostic versus fully specified model I



Note: The figure plots the distribution of χ^2 statistics of the first Monte Carlo experiment and the theoretical distribution according to large sample theory. This Monte Carlo experiment corresponds to the case when the true *dgp* does not include a monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

Figure 3.8: Likelihood ratio test agnostic versus fully specified model II



Note: The figure plots the distribution of χ^2 statistics of the first Monte Carlo experiment and the theoretical distribution according to large sample theory. This Monte Carlo experiment corresponds to the case when the true *dgp* does not include a TFP disturbance, but the empirical model leaves out the investment disturbance instead.

Table 3.9: Parameter estimates across Monte Carlo replications I

				misspecified estimation					ASD procedure					SW specification				
	Truth	**LB**	**UB**	**10%**	**25%**	**50%**	**75%**	**90%**	**10%**	**25%**	**50%**	**75%**	**90%**	**10%**	**25%**	**50%**	**75%**	**90%**
α	0.19	0.07	0.31	0.07	0.07	0.08	0.10	0.14	0.15	0.18	0.21	0.23	0.26	0.15	0.17	0.20	0.22	0.24
σ_c	1.39	0.53	2.25	1.35	1.44	1.66	2.16	2.24	1.16	1.29	1.47	1.69	1.95	1.19	1.29	1.42	1.58	1.81
Φ	1.61	1.33	1.89	1.46	1.69	1.86	1.89	1.89	1.33	1.34	1.57	1.85	1.89	1.33	1.38	1.60	1.83	1.89
ϕ	5.48	1.99	8.97	3.42	5.03	6.48	8.05	8.87	3.18	3.78	4.59	5.63	6.74	3.80	4.27	4.91	5.55	6.24
λ	0.71	0.45	0.90	0.62	0.71	0.79	0.84	0.86	0.57	0.62	0.66	0.71	0.76	0.61	0.64	0.68	0.72	0.76
ξ_w	0.73	0.47	0.92	0.58	0.65	0.71	0.78	0.86	0.57	0.62	0.68	0.75	0.82	0.58	0.63	0.69	0.75	0.81
σ_ℓ	1.92	0.18	3.66	0.18	0.18	0.18	0.19	0.54	0.89	1.16	1.55	2.17	2.77	1.07	1.36	1.76	2.24	2.71
ξ_p	0.65	0.40	0.86	0.69	0.74	0.79	0.83	0.85	0.50	0.56	0.63	0.69	0.73	0.52	0.57	0.63	0.68	0.73
ι_w	0.59	0.24	0.89	0.24	0.28	0.38	0.53	0.68	0.32	0.44	0.60	0.76	0.88	0.34	0.47	0.61	0.76	0.88
ι_p	0.22	0.01	0.65	0.04	0.13	0.25	0.38	0.51	0.02	0.08	0.16	0.25	0.34	0.03	0.08	0.16	0.24	0.31
ψ	0.54	0.20	0.86	0.26	0.38	0.59	0.75	0.84	0.38	0.47	0.57	0.69	0.81	0.41	0.48	0.57	0.66	0.77
r_π	2.03	1.45	2.61	1.58	1.78	2.05	2.33	2.55	1.71	1.88	2.10	2.41	2.60	1.76	1.91	2.09	2.35	2.57
ρ	0.81	0.53	0.97	0.74	0.77	0.80	0.82	0.83	0.78	0.80	0.82	0.84	0.85	0.78	0.80	0.82	0.84	0.85
r_y	0.08	-0.04	0.20	0.05	0.07	0.11	0.17	0.20	0.05	0.07	0.08	0.11	0.13	0.06	0.07	0.08	0.10	0.12
$r_{\Delta y}$	0.22	0.10	0.34	0.11	0.15	0.17	0.18	0.19	0.19	0.21	0.22	0.23	0.24	0.20	0.21	0.22	0.23	0.24
ρ_a	0.95	0.00	0.99	0.60	0.93	0.95	0.96	0.96	0.88	0.92	0.94	0.95	0.96	0.90	0.93	0.94	0.96	0.96
ρ_b	0.18	0.00	0.99	0.03	0.08	0.16	0.27	0.75	0.03	0.08	0.14	0.20	0.26	0.04	0.09	0.15	0.21	0.26
ρ_g	0.97	0.00	0.99	0.99	0.99	0.99	0.99	0.99	0.94	0.96	0.97	0.98	0.98	0.94	0.96	0.97	0.98	0.98
ρ_p	0.90	0.00	0.99	0.50	0.66	0.78	0.91	0.97	0.68	0.80	0.87	0.92	0.95	0.74	0.82	0.88	0.92	0.94
ρ_w	0.97	0.00	0.99	0.93	0.97	0.99	0.99	0.99	0.93	0.95	0.97	0.98	0.99	0.94	0.96	0.97	0.98	0.99
μ_p	0.74	0.00	0.99	0.21	0.38	0.64	0.86	0.94	0.31	0.48	0.62	0.73	0.80	0.38	0.51	0.64	0.73	0.80
μ_w	0.88	0.00	0.99	0.72	0.81	0.87	0.91	0.94	0.73	0.80	0.85	0.89	0.92	0.76	0.82	0.86	0.89	0.92
σ_a	0.45	0.00	10.00	0.62	0.70	0.85	1.05	1.22	0.35	0.38	0.42	0.48	0.53	0.37	0.39	0.44	0.48	0.53
σ_b	0.24	0.00	10.00	0.06	0.20	0.24	0.26	0.28	0.21	0.22	0.24	0.26	0.28	0.21	0.23	0.24	0.26	0.28
σ_g	0.52	0.00	10.00	0.51	0.53	0.55	0.57	0.59	0.48	0.49	0.52	0.54	0.56	0.48	0.50	0.52	0.54	0.56
σ_p	0.14	0.00	10.00	0.12	0.14	0.15	0.17	0.18	0.11	0.12	0.14	0.15	0.17	0.11	0.12	0.14	0.15	0.16
σ_w	0.24	0.00	10.00	0.19	0.20	0.22	0.24	0.25	0.21	0.23	0.24	0.26	0.28	0.21	0.23	0.25	0.26	0.28

Note: The table provides information on the distribution of the indicated parameter across the Monte Carlo replications. See Table 3.1 for the definitions of the parameters. This is for the first Monte Carlo experiment which corresponds to the case when the true *dgp* does not include a monetary policy disturbance, but the empirical model leaves out the investment disturbance instead.

Table 3.10: Parameter estimates across Monte Carlo replications II

				misspecified estimation					ASD procedure					SW specification				
	Truth	**LB**	**UB**	**10%**	**25%**	**50%**	**75%**	**90%**	**10%**	**25%**	**50%**	**75%**	**90%**	**10%**	**25%**	**50%**	**75%**	**90%**
α	0.19	0.07	0.31	0.09	0.12	0.16	0.21	0.26	0.14	0.18	0.21	0.25	0.29	0.14	0.17	0.20	0.23	0.26
σ_c	1.39	0.53	2.25	1.39	1.57	2.07	2.24	2.25	1.14	1.27	1.45	1.75	2.14	1.16	1.26	1.41	1.63	1.93
Φ	1.61	1.33	1.89	1.39	1.58	1.79	1.87	1.89	1.33	1.35	1.58	1.85	1.89	1.34	1.43	1.61	1.78	1.87
ϕ	5.48	1.99	8.97	4.09	4.99	6.23	7.43	8.48	3.29	3.85	4.57	5.43	6.54	3.85	4.34	4.98	5.64	6.44
λ	0.71	0.45	0.90	0.54	0.61	0.69	0.76	0.80	0.55	0.61	0.66	0.72	0.77	0.60	0.64	0.68	0.73	0.77
ξ_w	0.73	0.47	0.92	0.55	0.61	0.67	0.72	0.78	0.55	0.60	0.67	0.74	0.81	0.58	0.62	0.69	0.75	0.80
σ_ℓ	1.92	0.18	3.66	0.18	0.18	0.20	0.55	1.03	0.46	0.89	1.56	2.58	3.54	0.67	1.07	1.78	2.64	3.38
ξ_p	0.65	0.40	0.86	0.62	0.68	0.73	0.78	0.83	0.49	0.54	0.62	0.68	0.73	0.51	0.56	0.62	0.68	0.72
ι_w	0.59	0.24	0.89	0.24	0.30	0.42	0.57	0.69	0.32	0.46	0.61	0.78	0.89	0.34	0.47	0.62	0.77	0.88
ι_p	0.22	0.01	0.65	0.04	0.11	0.24	0.34	0.46	0.01	0.07	0.15	0.24	0.32	0.03	0.08	0.16	0.24	0.32
ψ	0.54	0.20	0.86	0.30	0.40	0.53	0.67	0.80	0.37	0.45	0.57	0.70	0.81	0.40	0.48	0.58	0.68	0.79
r_π	2.03	1.45	2.61	1.74	2.04	2.32	2.54	2.60	1.59	1.86	2.28	2.59	2.61	1.68	1.89	2.22	2.52	2.60
ρ	0.81	0.53	0.97	0.79	0.81	0.84	0.86	0.87	0.73	0.77	0.81	0.84	0.86	0.76	0.78	0.81	0.84	0.85
r_y	0.08	-0.04	0.20	0.08	0.10	0.13	0.16	0.19	0.04	0.06	0.10	0.12	0.14	0.05	0.07	0.09	0.12	0.13
$r_{\Delta y}$	0.22	0.10	0.34	0.11	0.13	0.17	0.20	0.22	0.17	0.19	0.23	0.26	0.29	0.17	0.20	0.22	0.25	0.28
ρ_b	0.18	0.00	0.99	0.07	0.14	0.22	0.33	0.53	0.02	0.08	0.14	0.21	0.26	0.04	0.09	0.15	0.21	0.26
ρ_g	0.97	0.00	0.99	0.98	0.99	0.99	0.99	0.99	0.94	0.96	0.97	0.98	0.98	0.94	0.96	0.97	0.98	0.98
ρ_r	0.12	0.00	0.99	0.00	0.00	0.02	0.06	0.10	0.00	0.04	0.11	0.18	0.23	0.01	0.05	0.11	0.17	0.22
ρ_p	0.90	0.00	0.99	0.48	0.65	0.78	0.91	0.97	0.74	0.82	0.89	0.92	0.95	0.78	0.84	0.89	0.92	0.95
ρ_w	0.97	0.00	0.99	0.94	0.96	0.97	0.98	0.99	0.93	0.95	0.97	0.98	0.99	0.94	0.96	0.97	0.98	0.98
μ_p	0.74	0.00	0.99	0.17	0.37	0.61	0.84	0.92	0.36	0.50	0.62	0.72	0.81	0.44	0.55	0.65	0.73	0.80
μ_w	0.88	0.00	0.99	0.73	0.79	0.84	0.89	0.91	0.73	0.79	0.84	0.89	0.91	0.76	0.81	0.85	0.89	0.92
σ_b	0.24	0.00	10.00	0.14	0.19	0.22	0.24	0.26	0.20	0.22	0.24	0.26	0.28	0.21	0.23	0.24	0.26	0.28
σ_g	0.52	0.00	10.00	0.49	0.51	0.53	0.55	0.57	0.47	0.49	0.51	0.54	0.56	0.48	0.50	0.52	0.54	0.56
σ_r	0.24	0.00	10.00	0.22	0.23	0.24	0.25	0.26	0.22	0.23	0.24	0.25	0.26	0.22	0.23	0.24	0.25	0.26
σ_p	0.14	0.00	10.00	0.12	0.14	0.15	0.16	0.18	0.11	0.12	0.14	0.15	0.17	0.11	0.12	0.14	0.15	0.16
σ_w	0.24	0.00	10.00	0.20	0.21	0.23	0.24	0.26	0.21	0.23	0.25	0.27	0.28	0.22	0.23	0.25	0.27	0.28

Note: The table provides information on the distribution of the indicated parameter across the Monte Carlo replications. See Table 3.1 for the definitions of the parameters. This for the second Monte Carlo experiment which corresponds to the case when the true *dgp* does not include a TFP disturbance, but the empirical model leaves out the investment disturbance instead.

3.7.4 ASD procedure for the Smets-Wouters model

In this appendix, we provide further details on how the ASD procedure is implemented in section 3.6. We also provide additional results.

Including ASDs in SW equations

To apply the ASD procedure to the SW model, we adapt the Dynare program provided by the authors.[77] Adapting a Dynare program to add an agnostic disturbance is easy. Specifically, for the first ASD, $\widetilde{\varepsilon}_t{}^A$, we do the following.

1. In the model block, we add $\boldsymbol{d}_i^A \widetilde{\varepsilon}_t^A$ to the i^{th} equation, where $\widetilde{\varepsilon}_t^A$ is the agnostic disturbance and $\boldsymbol{d}_i^A$ the coefficient associated with the agnostic disturbance in the i^{th} equation. Details are given below.[78]

2. We add an equation to the model block that describes the law of motion for $\widetilde{\varepsilon}_t^A$. If the agnostic disturbance replaces a regular structural disturbance, then this disturbance should be taken out of the program.

3. Declare $\widetilde{\varepsilon}_t^A$ as a variable and declare the elements of $\boldsymbol{d}_i^A$ and the coefficients of the law of motion for $\widetilde{\varepsilon}_t^A$ as parameters.

4. Specify a prior for the elements of $\boldsymbol{d}_i^A$.

We do not add the agnostic disturbance to equations (6) and (12) of the SW model, because these equations just contain definitions for capacity utilization and the wage mark-up, respectively.[79] The set of equations for the SW model consists of two parts. The first part models the flexible price economy and the second part models the actual economy with sticky prices. One needs to model the flexible-price economy, because the flexible-price output level is used to define the output gap, which is one of the arguments in the monetary policy rule. In principle, one could let the agnostic disturbance enter the equations of the sticky-price economy and the associated equations in the flexible-price economy with a different coefficient.[80] In most economic models, however, structural disturbances would enter the associated pair of equations in the same way. Therefore, we also restrict the agnostic disturbance to enter the associated equations in the same way. The exception is SW equation (13) because it captures both potential stickiness in wages and the relationship between the wage rate and its mark-up.

[77] The program is available at https://www.aeaweb.org/articles?id=10.1257/aer.97.3.586 under the "Download Data Set" link.

[78] The other two ASDs are added using the same procedure. The $\boldsymbol{d}^A$ coefficients correspond to the $\widehat{\boldsymbol{\Gamma}}_{\mathbf{2}}$ coefficients in section 3.3.2. We adapt the notation, since SW also use lower case Roman letters for coefficients.

[79] Equation numbers refer to those in Smets and Wouters (2007). We do allow the agnostic disturbances to affect the utilization rate and the wage mark-up directly by including it in the model equations that specify their relationship with other model variables.

[80] The sticky-price block contains some equations, such as the monetary policy rule, that do not have a counterpart in the flexible-price economy.

Specifically, we add the agnostic disturbance to equations (1), (2), (3), (4), (5), (7), (8), (9), and (11) of the SW model and the associated equations of the flexible-price economy. We also add it to equation (13) in both the flexible and the sticky-price part of the model, but here we allow coefficients to differ. In addition, we add the agnostic disturbance to equations (10) and (14) which do not have a counterpart in the flexible-price economy. This means that $\boldsymbol{d}^A$ has thirteen elements. The last coefficient associated with the agnostic disturbance is the autoregressive coefficient of its law of motion. The standard deviation of the agnostic disturbance is normalized to be equal to 1.

Model selection procedures

The general-to-specific model selection procedure starts with the specification in which the agnostic disturbances are allowed to enter each model equation. It then calculates the marginal data densities for all possible specifications in which the ASD is *not* allowed to enter *one* of the model equations. Thus, we estimate a set of models, each having one less coefficient. If none of the specifications lead to a better fit, then the procedure stops. If improvements are found, then the procedure is repeated using the specification that led to the biggest improvement as the benchmark.

The specific-to-general procedure starts with the specifications in which each of the two ASDs are allowed to enter only one model equation. To avoid a singularity, one cannot start with a more parsimonious model.[81] In the next step, we estimate a set of models in which one of the ASDs is added to one equation and, thus, one additional parameter is estimated. The procedure stops if none of the specifications leads to an improvement. If there is an improvement, then the specification with the largest improvement becomes the next benchmark and the procedure is repeated.

Why not consider even more general specifications? Although our model selection procedures consider a rich set of models, they are not the most general. Unfortunately, there are practical limitations to what is feasible. Five SW disturbance are always included in our specifications. The most ideal setup would be flexible in this dimension as well and not safeguard any of the seven SW regular disturbances and allow for the possibility of including seven ASDs (or more). With such a setup all SW disturbances could be replaced by an ASD. The first problem one would have to deal with is that identification of structural parameters is likely to limit the number of regular structural disturbances one can replace with ASDs. Let us consider a simple setup in which there are seven equations for seven state variables and all state variables are observables. Moreover, each equation has one regular structural disturbance. A general-to-specific procedure would be complicated since the first-stage model would have a large number of coefficients to estimate. Specifically, if all seven ASDs appear in all equations, then one needs to estimate forty-nine reduced-form

[81]The posteriors of the ASD coefficients in the fully agnostic model provide clear evidence that one of the ASDs is very important for the bond Euler equation and one for the investment Euler equation. So these are natural choices.

coefficients. One may need a rich data set to identify all of them. In our application, the number of coefficients would be equal to ninety-one, since we have thirteen equations. The specific-to-general procedure faces the problem that each specification needs at least seven disturbances to avoid singularities. This means that there are a large number of different models one can start with. For the simple setup with seven equations described above, this would mean that there are already $2^7 = 128$ different models to consider in the first round alone.

Additional results

Specifications with and without restrictions on ASDs. Table 3.11 compares structural parameter estimates of models chosen by our model selection procedures with those that contain the same number of ASDs, but allow ASDs to enter all equations. The latter are fully agnostic. The parameter estimates are fairly similar. IRFs for the included regular structural disturbances are also quite similar. That is not always the case for the IRFs of the agnostic disturbances themselves. The IRFs for some variables do differ between the concise and the fully unrestricted ASD specification. Given the misspecification results of section 3.2, it is not surprising that different empirical specifications lead to different results. Another issue with the fully unrestricted ASD specification is that it estimates a large number of coefficients which complicates generating an accurate posterior with Monte Carlo Markov Chain algorithms. Especially, for the 3-ASD fully unrestricted specification, the Brooks-Gelman statistics did not look particularly good for some of the coefficients associated with the agnostic disturbances.

Specifications with two and three ASDs. Tables 3.12 and 3.13 provide the role of the regular and agnostic disturbances for the fluctuations of a wide range of variables. In addition to the results of the SW specification, it also shows the results for the two-ASD and three-ASD specification chosen by our specific-to-general procedure. It shows that the results are very similar for the two chosen ASD specifications. The same conclusion can be drawn from Figures 3.9 and 3.10 that plot the IRFs for two agnostic disturbances.

Table 3.11: Posterior Means

Parameter	Original SW	Agnostic: 2 ASDs		Agnostic: 3 ASDs	
		concise	unrestricted	concise	unrestricted
α	0.1903	0.2044	0.1878	0.1877	0.2089
σ_c	1.3889	1.4657	1.4535	1.4618	1.4772
Φ	1.6083	1.5211	1.5242	1.4741	1.4762
ϕ	5.7405	5.3843	4.4031	5.3425	4.6933
λ	0.7136	0.6544	0.7055	0.6679	0.6930
ξ_w	0.7066	0.6660	0.6706	0.7268	0.6453
σ_ℓ	1.8458	1.9094	1.7733	2.0770	1.5916
ξ_p	0.6541	0.6566	0.6981	0.6412	0.6902
ι_w	0.5783	0.5556	0.5432	0.5077	0.5557
ι_p	0.2389	0.2010	0.1997	0.1871	0.1891
ψ	0.5426	0.5345	0.5049	0.5283	0.3176
r_π	2.0469	1.7676	1.7797	1.7746	1.7438
ρ	0.8105	0.7933	0.8082	0.8018	0.8032
r_y	0.0887	0.0725	0.0860	0.0787	0.0819
$r_{\Delta y}$	0.2237	0.1903	0.1703	0.1941	0.1608
ρ_a	0.9572	0.9555	0.9483	0.9532	0.9510
ρ_g	0.9764	0.9719	0.9710	0.9702	0.9018
ρ_r	0.1464	0.1376	0.1219	0.1286	0.1227
ρ_p	0.8893	0.8975	0.8899	0.9262	0.9080
ρ_w	0.9680	0.9751	0.9790	0.9747	0.9822
ρ_b / ρ_A	0.2165	0.3344	0.6386	0.3239	0.4527
ρ_i / ρ_B	0.7116	0.6087	0.1660	0.6069	0.7232
ρ_C	-	-	-	0.1865	0.1577
μ_p	0.6977	0.6764	0.6923	0.7166	0.7172
μ_w	0.8466	0.8241	0.8368	0.5945	0.8168
ρ_{ga}	0.5184	0.6438	0.6525	0.6709	0.5448
σ_a	0.4586	0.4436	0.4421	0.4524	0.4411
σ_g	0.5299	0.4702	0.4689	0.4428	0.2285
σ_r	0.2449	0.2180	0.2171	0.2171	0.2114
σ_p	0.1403	0.1346	0.1299	0.1308	0.1311
σ_w	0.2427	0.2384	0.2361	0.0763	0.2249
σ_b	0.2398	-	-	-	-
σ_i	0.4525	-	-	-	-
$100(\beta^{-1}-1)$	0.1648	0.1685	0.1826	0.1656	0.2038
$\bar{\gamma}$	0.4316	0.4349	0.4386	0.4367	0.4352
$\bar{\pi}$	0.7845	0.7483	0.7443	0.7391	0.7534
$\bar{\ell}$	0.5617	0.1263	0.5216	0.1303	1.0360
MDD	-922.40	-892.92	-906.85	-890.73	-925.50

Note: MDD stands for marginal data density. The “concise" ASD specifications are the ones chosen by the specific-to-general model selection procedure. The “unrestricted" ASD specifications are the fully agnostic with no zero restrictions. See Table 3.1 for the definitions of the parameters.

Table 3.12: VARIANCE DECOMPOSITIONS ACROSS MODEL SPECIFICATIONS

		ε^a	ε^g	ε^r	ε^p	ε^w	$\varepsilon^b/\widetilde{\varepsilon}^A$	$\varepsilon^i/\widetilde{\varepsilon}^B$	$\widetilde{\varepsilon}^C$
Δy	Original SW	16.10	28.88	6.17	4.55	6.39	22.12	15.79	-
	Agnostic: 2 ASDs	20.29	27.01	7.15	6.04	8.12	20.53	10.85	-
	Agnostic: 3 ASDs	22.21	24.60	7.04	4.66	10.30	21.33	8.04	1.82
Δc	Original SW	5.29	2.10	11.56	4.40	14.54	61.17	0.95	-
	Agnostic: 2 ASDs	3.26	1.62	11.29	4.56	15.33	62.34	1.61	-
	Agnostic: 3 ASDs	2.95	1.28	10.69	3.37	17.90	61.67	2.03	0.1
Δi	Original SW	6.01	0.84	2.47	3.80	2.37	2.46	82.05	-
	Agnostic: 2 ASDs	4.86	0.91	2.19	4.24	2.76	12.25	72.80	-
	Agnostic: 3 ASDs	5.49	1.02	2.38	3.80	3.94	12.55	70.01	0.81
ℓ	Original SW	1.94	10.34	3.15	6.23	67.66	2.52	8.15	-
	Agnostic: 2 ASDs	1.29	6.84	2.47	6.04	71.23	1.56	10.57	-
	Agnostic: 3 ASDs	1.08	4.33	2.15	4.44	79.70	1.29	4.97	2.03
Δw	Original SW	4.53	0.09	1.48	29.47	61.61	0.79	2.03	-
	Agnostic: 2 ASDs	3.82	0.22	2.43	30.84	54.34	3.02	5.34	-
	Agnostic: 3 ASDs	4.09	0.11	1.25	25.18	13.32	2.23	0.38	53.45
π	Original SW	3.92	1.00	4.25	27.64	59.43	0.58	3.18	-
	Agnostic: 2 ASDs	3.16	1.28	4.43	24.91	61.96	0.79	3.46	-
	Agnostic: 3 ASDs	2.95	0.90	3.28	16.87	70.46	0.68	3.96	0.91
r	Original SW	10.09	3.90	14.67	7.17	38.42	7.40	18.34	-
	Agnostic: 2 ASDs	6.50	3.49	9.77	5.79	38.96	21.49	14.02	-
	Agnostic: 3 ASDs	5.70	2.77	8.18	4.33	48.61	17.29	12.47	0.65

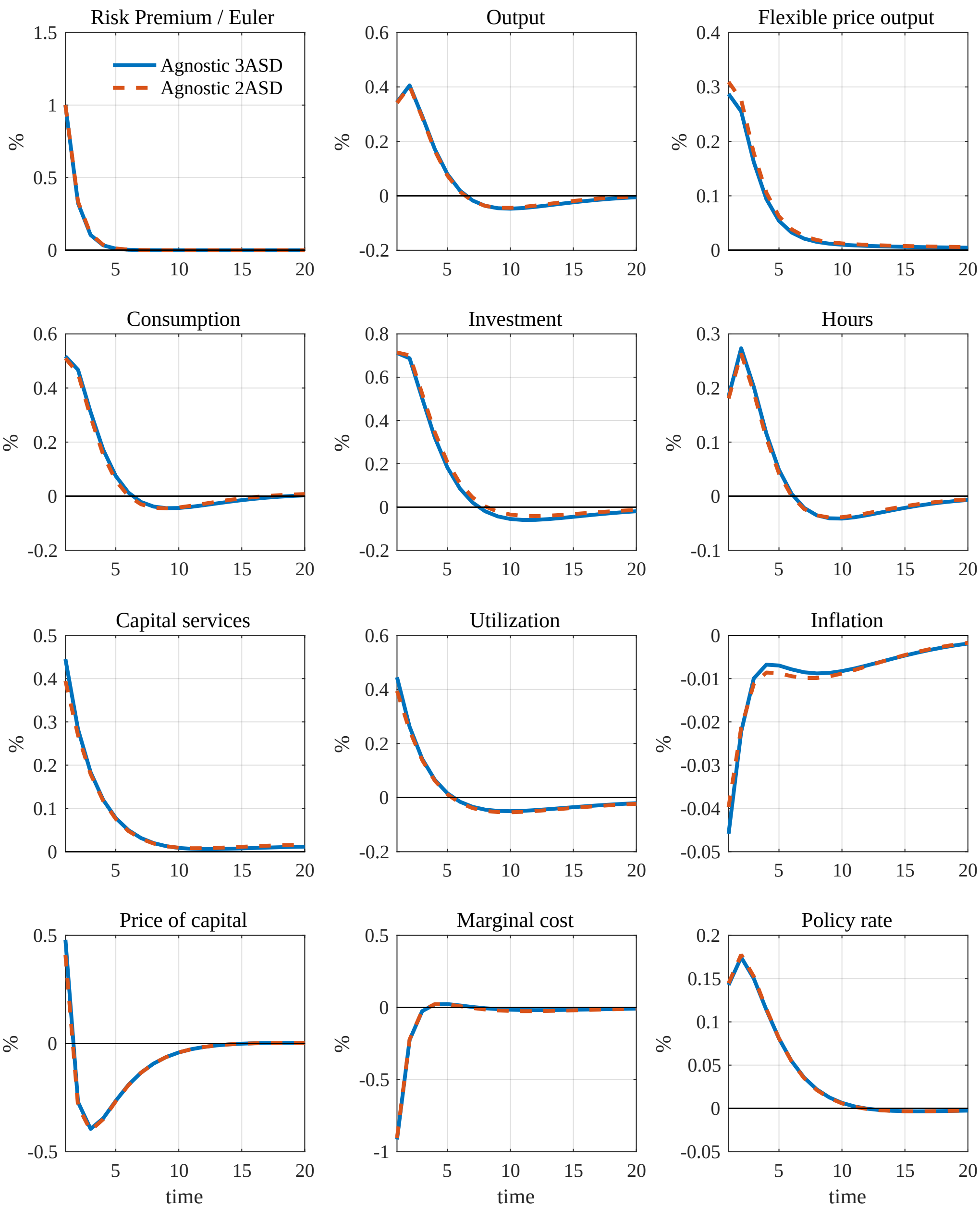
Note: The table provides the contributions (in percent) of the different structural disturbances to the variance of the observable variables, across different model specifications. The ASD specifications are the ones chosen by our model selection procedure. y stands for log output; c for log consumption; i for log investment; l for hours; w for log wage rate; π for inflation; and r for nominal interest rate. Structural disturbances are defined as follows. ε^a: TFP; ε^g: government expenditures; ε^r: monetary policy; ε^p:price mark-up; ε^w: wage mark-up; ε^b: risk premium; ε^i: investment; $\widetilde{\varepsilon}^A$: agnostic Euler; $\widetilde{\varepsilon}^B$: agnostic investment-modernization; and $\widetilde{\varepsilon}^C$: capital-efficiency wage mark-up.

Table 3.13: Variance decomposition for additional variables

		ε^a	ε^g	ε^r	ε^p	ε^w	$\varepsilon^b/\widetilde{\varepsilon}^A$	$\varepsilon^i/\widetilde{\varepsilon}^B$	$\widetilde{\varepsilon}^C$
y_t	Original SW	29.93	4.09	2.16	6.37	48.58	1.53	7.34	-
	Agnostic: 2 ASDs	26.50	3.02	1.91	7.02	55.93	1.31	4.32	-
	Agnostic: 3 ASDs	21.19	2.13	1.67	5.47	65.95	1.14	2.17	0.28
c_t	Original SW	11.06	8.42	2.08	4.19	69.25	2.18	2.83	-
	Agnostic: 2 ASDs	6.60	6.60	1.78	4.23	78.76	1.81	0.22	-
	Agnostic: 3 ASDs	4.29	4.30	1.52	3.16	84.48	1.51	0.49	0.25
i_t	Original SW	20.37	5.41	1.27	6.93	21.56	0.22	44.23	-
	Agnostic: 2 ASDs	17.22	6.35	1.14	8.25	29.78	1.21	36.04	-
	Agnostic: 3 ASDs	15.31	5.79	1.13	7.75	38.72	1.06	29.25	1.00
r_t^k	Original SW	14.86	17.47	1.63	10.58	19.21	0.86	35.39	-
	Agnostic: 2 ASDs	12.28	20.44	2.65	19.09	29.73	0.92	14.88	-
	Agnostic: 3 ASDs	8.41	14.66	1.73	13.16	30.17	0.67	18.12	13.08
q_t	Original SW	4.65	0.55	9.03	3.11	1.20	45.42	36.04	-
	Agnostic: 2 ASDs	9.78	1.35	21.83	9.30	3.67	19.58	34.49	-
	Agnostic: 3 ASDs	9.72	1.35	19.88	6.50	5.14	18.64	31.56	7.21
z_t	Original SW	14.86	17.47	1.63	10.58	19.21	0.86	35.39	-
	Agnostic: 2 ASDs	12.23	20.36	2.64	19.02	29.61	4.43	11.71	-
	Agnostic: 3 ASDs	8.86	15.43	1.82	13.85	31.76	4.14	9.46	14.68
μ_t^p	Original SW	11.56	0.29	3.27	57.02	23.87	0.87	3.11	-
	Agnostic: 2 ASDs	8.06	0.37	3.38	53.22	18.90	14.59	1.48	-
	Agnostic: 3 ASDs	7.99	0.24	2.13	54.80	11.88	15.22	2.61	5.13
k_t^s	Original SW	23.43	3.92	1.23	11.37	34.19	0.36	25.50	-
	Agnostic: 2 ASDs	21.82	4.90	1.55	16.51	52.66	1.32	1.24	-
	Agnostic: 3 ASDs	15.59	3.60	1.11	14.11	58.20	1.21	0.61	5.57
k_t	Original SW	22.38	8.11	0.50	4.93	31.56	0.04	32.48	-
	Agnostic: 2 ASDs	22.16	11.30	0.55	7.27	55.74	0.21	2.77	-
	Agnostic: 3 ASDs	14.84	8.05	0.42	6.18	58.26	0.12	2.37	9.75
w_t	Original SW	33.03	1.03	1.95	38.38	18.61	0.40	6.60	-
	Agnostic: 2 ASDs	26.99	1.00	2.63	47.34	20.71	0.39	0.92	-
	Agnostic: 3 ASDs	25.35	0.74	1.62	49.34	14.29	0.30	0.44	7.92

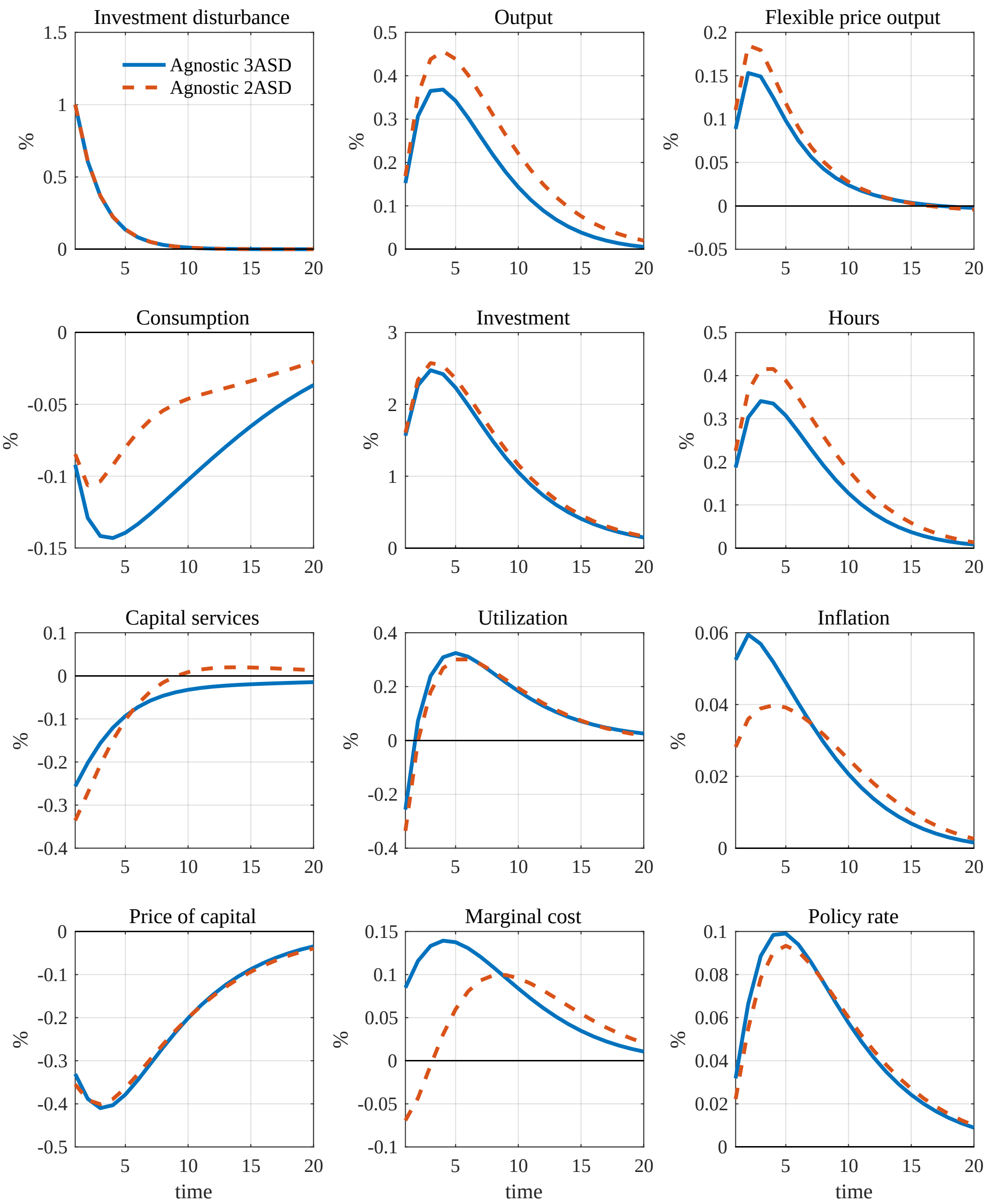
Note: The table provides the contributions (in percent) of the different structural disturbances to the variance of the observable variables, across different model specifications. The ASD specifications are the ones chosen by our model selection procedure. y stands for log output; c for log consumption; i for log investment; l for hours; w for log wage rate; r^k for rental rate on capital; q for the log price of capital; z for the utilization rate; μ^p for the price mark-up; k^s for log capital used in production; and k for log installed capital. Structural disturbances are defined as follows. ε^a: TFP; ε^g: government expenditures; ε^r: monetary policy; ε^p:price mark-up; ε^w: wage mark-up; ε^b: risk premium; ε^i: investment; $\widetilde{\varepsilon}^A$: agnostic Euler; $\widetilde{\varepsilon}^B$: agnostic investment-modernization; and $\widetilde{\varepsilon}^C$: capital-efficiency wage mark-up.

Figure 3.9: IRFs of the Euler ASD: 2 versus 3 ASDs



Note: These figures plot the IRFs of the agnostic disturbance $\widetilde{\varepsilon}_t^A$ that we interpret as a general Euler disturbance for the empirical specifications with two and three ASDs. Both are chosen with the specific-to-general model selection procedure.

Figure 3.10: IRFs of the investment-modernization ASD: 2 versus 3 ASDs



Note: These figures plot the IRFs of the agnostic disturbance $\tilde{\varepsilon}_t^B$ that we interpret as an investment-modernization disturbance for the empirical specifications with two and three ASDs. Both are chosen with the specific-to-general model selection procedure.

Additional results for $\widetilde{\varepsilon}_t^A$

Figure 3.11 plots the IRFs associated with an innovation in the agnostic Euler disturbance for our 3-ASD benchmark specification and also when the coefficient of this agnostic disturbance in the capital valuation equation is equal to zero. A zero coefficient in this equation means the disturbance is like a preference and not like a bond risk-premium disturbance.[82] The IRFs are very similar, which confirms our claim that the coefficient in the capital valuation equation is quantitatively not very important.

Figure 3.12 plots the same IRFs when the coefficient of the agnostic Euler disturbance in the Taylor rule is set equal to zero. The figure shows that the direct response of the policy rate to a positive shock to this disturbance dampens the expansion and prevents an upsurge of inflation.

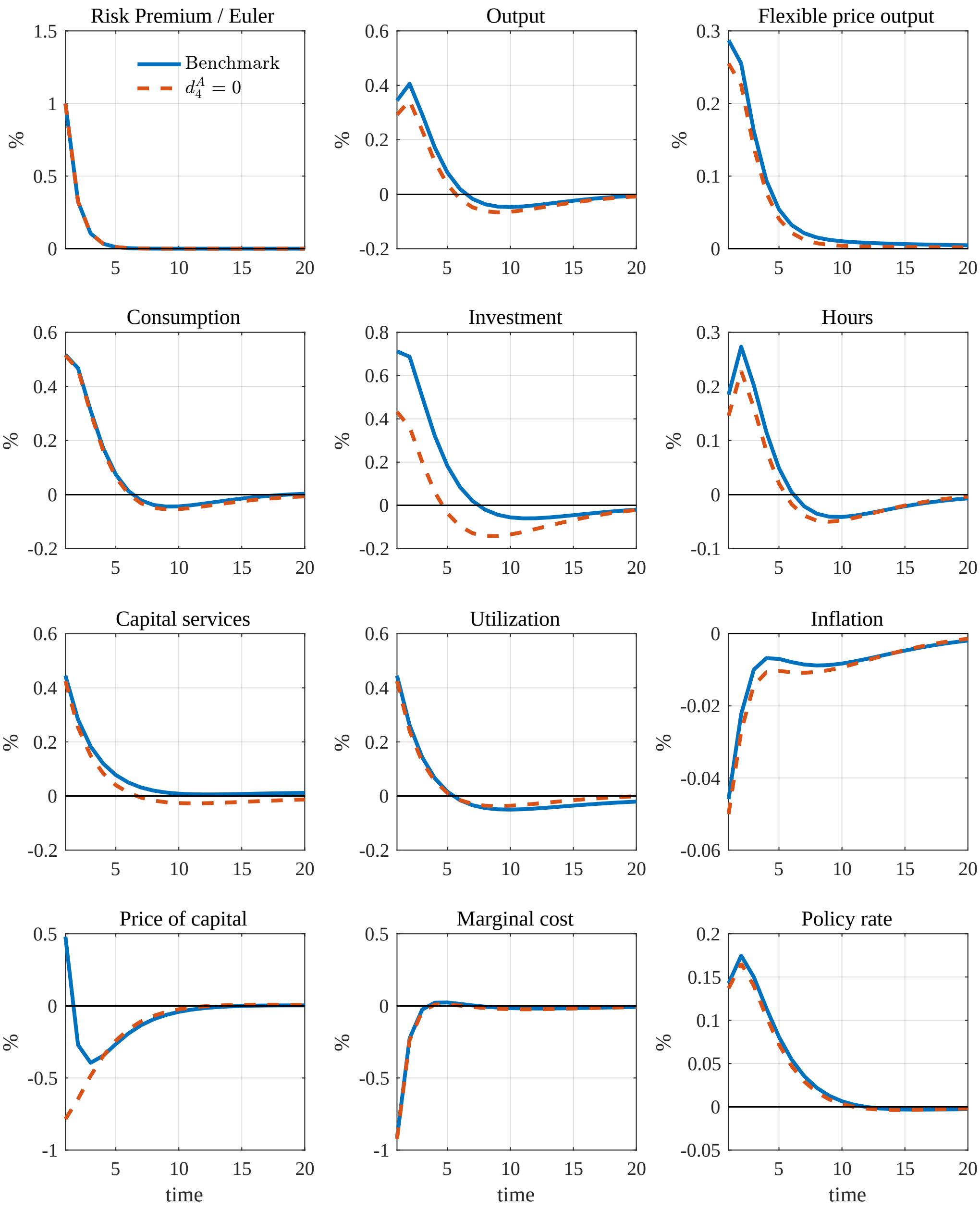
Figure 3.13 plots the same IRFs when we set equal to zero the coefficients of the disturbance in the four equations that we ignored in the discussion of the agnostic Euler disturbance, namely, the overall budget constraint, the utilization, the price mark-up equation, and the rental rate of capital equation. The figure documents that the role of the agnostic disturbance through these equations is minor since the IRFs are overall quite similar to those of our benchmark specification.

[82]Recall that the MRS has been substituted out of the capital valuation equation using the MRS of the bond Euler equation.

Additional results for $\widetilde{\varepsilon}_t^C$

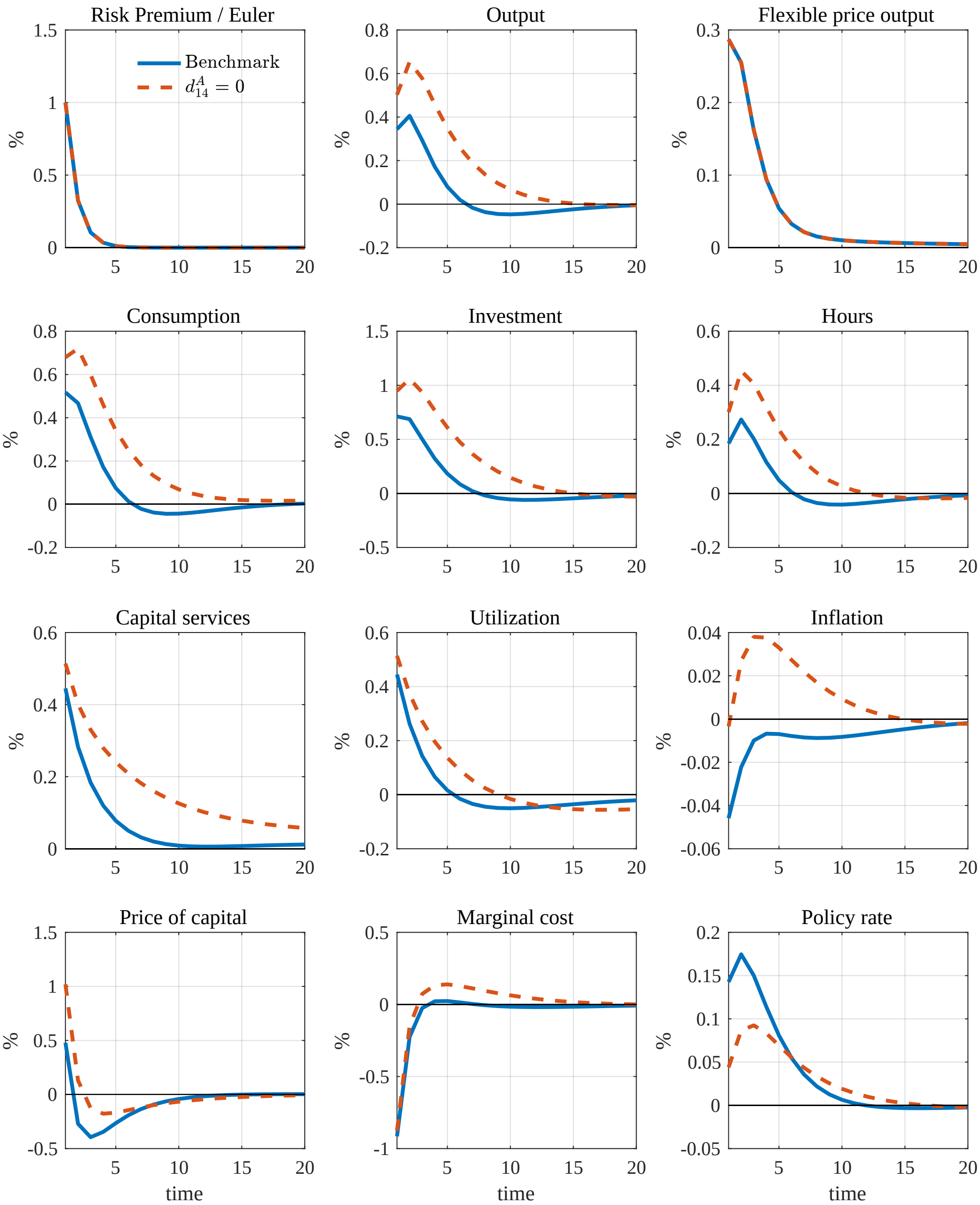
Figure 3.14 plots the IRFs for our agnostic capital-efficiency wage mark-up disturbance when the coefficient of this disturbance in the overall budget constraint is set equal to zero. The figure documents that this has a minor impact on IRFs.

Figure 3.11: IRFs of the Euler ASD with restrictions I



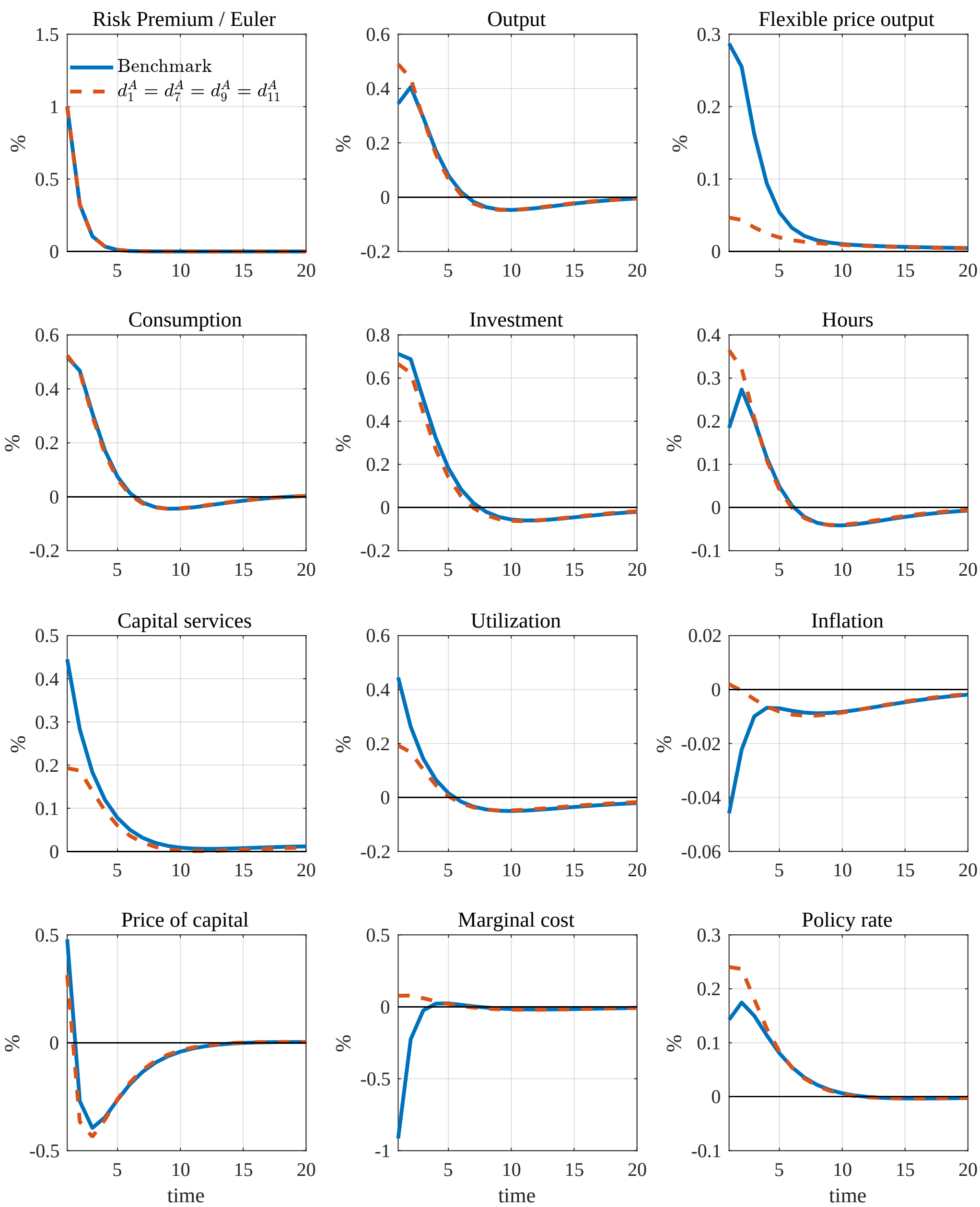
Note: These figures plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the capital valuation equation is set equal to zero.

Figure 3.12: IRFs of the Euler ASD with restrictions II



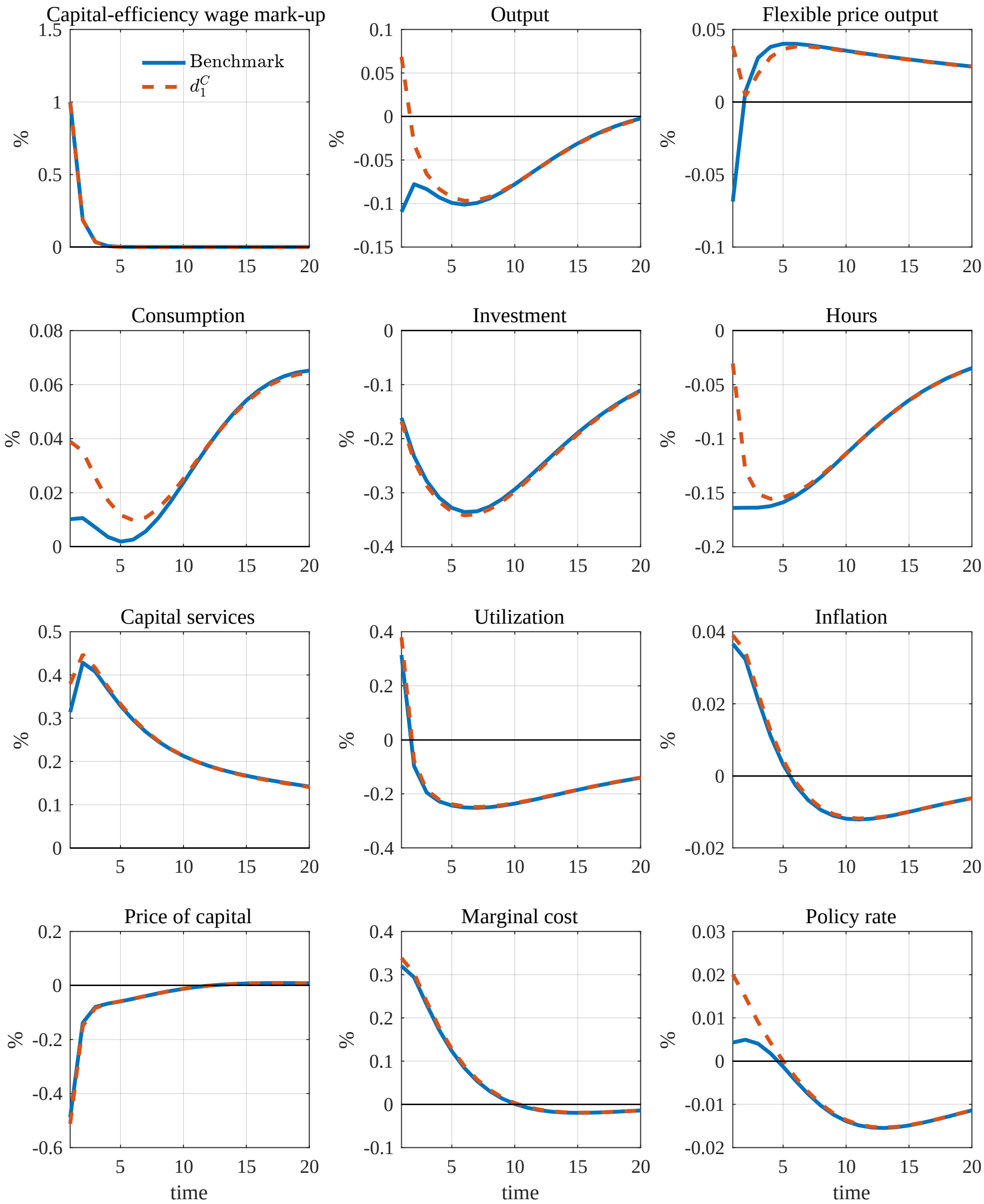
Note: These figures plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the Taylor rule is set equal to zero.

Figure 3.13: IRFs of the Euler ASD with restrictions III



Note: These figures plot the IRFs of the agnostic Euler disturbance for our benchmark specification and when the impact of this IRF through the overall budget constraint, the utilization, the price mark-up equation, and the rental rate of capital equation is set equal to zero.

Figure 3.14: IRFs of the capital-efficiency wage mark-up ASD with restrictions



Note: These figures plot the IRFs of the agnostic capital-efficiency wage mark-up disturbance for our benchmark specification and when the impact of this IRF through the overall budget constraint is set equal to zero.

3.7.5 Misspecification: Literature review

Most empirical papers that estimate a dynamic macroeconomic model do not raise the issue of model uncertainty or misspecification, except possibly with some robustness exercises.[83] This does – of course – not mean that the profession is not aware that misspecification is a serious concern. In fact, some of the most prominent researchers in this research area have drawn attention to the risk of misspecification. The first subsection discusses evidence that indicates that misspecification of DSGE models is a serious concern. The second subsection discusses approaches proposed in the literature to deal with misspecfication. See Paccagnini (2017) for a more detailed survey.

Indications of DSGE misspecification

Del Negro et al. (2007) develop a procedure that allows the data to determine the usefulness of a DSGE model relative to a much less restricted VAR. Using a model very similar to the DSGE model of Smets and Wouters (2003), they find that their procedure does put some weight on the DSGE model, which implies that the restrictions of the DSGE model are of some value. However, they also argue that misspecification is a concern that *"... is not small enough to be ignored."* Using the same methodology, Del Negro and Schorfheide (2009) also find *"... strong evidence of DSGE model misspecification."*

There is also more indirect evidence that misspecification of estimated DSGE models is substantive. Using the Smets and Wouters (2003) model for the Euro Area, Beltran and Draper (2015) find that the data prefer implausible estimates for several parameters. For example, *most* of the mass of the marginal likelihood for the parameter of relative risk aversion is above 200, way above the range of values considered reasonable. This information provided by the likelihood is typically not revealed in empirical studies, since only properties of the posterior are reported and the choice of prior ensures that these aspects of the empirical likelihood have little or no weight in the posterior. A similar conclusion can be drawn from Onatski and Williams (2010). They estimate the same model using uniform priors over bounded ranges. These ranges are such that the priors are less informative than the ones typically used in the literature. Consistent with the results in Beltran and Draper (2015), several of the point estimates in Onatski and Williams (2010) are at the prior bounds. Using a new algorithm to deal with the complexity of estimating likelihood functions, Mickelsson (2015) re-estimates the model of Smets and Wouters (2007) and he also finds that several parameter estimates are significantly different from the ones reported in Smets and Wouters (2007).

Another possible reason for misspecification is the assumption that parameters are constant. To get efficient estimates we would like to use long time-series data,

[83] Interestingly, there are quite a few macroeconomic models in which agents – especially agents setting fiscal and monetary policy – face model uncertainty. If policy makers face model uncertainty, then researchers are likely to do so as well.

but the longer the time series the less likely that all parameters are constant. Canova et al. (2015) address this issue and document that this is important for the model of Gertler and Karadi (2010).[84]

Dealing with misspecification: Other approaches

Richer models. Exogenous random disturbances are typically assumed not to be correlated with each other. This is a convenient assumption, because allowing for interaction between the different exogenous disturbances would substantially increase the number of parameters to be estimated given that DSGE typically have a several exogenous disturbances. However, it seems quite plausible that such disturbances are correlated. Del Negro and Schorfheide (2009) and Cúrdia and Reis (2012) deal with this possible misspecification and allow for more general processes to describe the behavior of the exogenous random disturbances.

Cúrdia and Reis (2012) find that this generalization has nontrivial consequences for the properties of the model. For example, the impact of a monetary policy shock on output is only half as big when the exogenous random variables are allowed to be correlated and the medium-term impact of a government spending shock switches from being positive to negative.[85]

Enriching a model by allowing for additional features and more general specifications is likely to reduce misspecification. However, richer models typically have more parameters, which will reduce the efficiency of the estimation by reducing the number of degrees of freedom.

Multiple models. Another way to deal with potential misspecification is to consider a set of different DSGE models. These could be compared informally or formally using, for example, relative marginal likelihoods or model averaging.[86] However, given the difficulty of modeling macroeconomic phenomena, it seems likely that *all* models in a set of DSGE models are subject to at least some type of misspecification.

Combining structural and reduced-form models. Ireland (2004) is an early paper that proposes a more general procedure to deal with possible misspecification when estimating a DSGE model even though the word misspecification is not used in the paper. Specifically, Ireland (2004) *"... augments the DSGE model so that its residuals – meaning the movements in the data that the theory cannot explain – are described by a VAR."* To understand this procedure, consider the following representation of the linearized solution of a DSGE model:

$$s_t = As_{t-1} + B\eta_t, \tag{3.57}$$

$$y_t = Cs_{t-1} + D\eta_t, \tag{3.58}$$

[84]The literature cited in Canova et al. (2015) documents that this is an issue in a variety of DSGE models.

[85]Cúrdia and Reis (2012) still impose that the *innovations* of the shocks are uncorrelated. Thus, the innovations still have a structural interpretation.

[86]See chapter 5 in An and Schorfheide (2007) for a detailed discussion.

where s_t is a vector containing (endogenous and exogenous) state variables, y_t is a vector containing the observables, and η_t is a vector containing the innovations of the exogenous random variables. Ireland (2004) proposes to augment the observation equation (3.58) as follows:

$$y_t = Cs_{t-1} + D\eta_t + u_t \tag{3.59a}$$

$$u_t = Fu_{t-1} + \xi_t \tag{3.59b}$$

where u_t captures the misspecification or incompleteness of the DSGE model. In his application, the structural equations are the policy rules from a standard Real Business Cycle (RBC) model with total factor productivity (TFP) as the only driving process. If the standard deviation of η_t is equal to 0, then this procedure boils down to estimating a standard VAR.

Note that the presence of the "missing elements" that are captured by u_t is assumed to have no effect on that part of agents' behavior that is described by the DSGE model, that is, the matrices A, B, C, and D. For this to be correct it must be true that the response of the economy to a TFP shock does not depend on the presence of other disturbances. One might think that such independence of a DSGE's policy rule to the presence of other disturbances is only correct if the additional disturbances represent measurement error.[87] However, section 3.3.2 shows that this "independence" property is correct in linear(ized) models in the sense that the specification of the structural part given in equations (3.57) and (3.58) does not depend on the presence of not included structural disturbances. It must be noted that the assumption that u_t follows a first-order (or even a finite-order) VAR could very well be restrictive. Thus the reduced-form specification for u_t could be misspecified as well.

The most comprehensive methodology to deal with misspecified DSGE models is put forward in Del Negro et al. (2007). Their starting point is a VAR specification of the observables. That is,

$$y_t = \sum_{k=1}^{K} F_k y_{t-1} + G\xi_t \tag{3.60a}$$

$$\mathbb{E}\left[\xi_t \xi_t'\right] = I. \tag{3.60b}$$

The key idea of the DSGE-VAR estimation proposed in Del Negro et al. (2007) is to estimate this time series process with the prior distribution for F and Ω that is centered at the values implied by a DSGE model, $F(\Psi)$ and $G(\Psi)$, where Ψ is the vector containing the parameters of the DSGE model. The estimation procedure consists of jointly estimating Ψ, the structural parameters of the DSGE model, which pin down the prior for the VAR coefficients, and the VAR coefficients themselves.

The precision of the prior of the VAR coefficients is controlled with a scalar

[87] Although Ireland (2004) does not refer to the residual between model and data as measurement error, other papers in the literature describing his procedure do. Examples are Del Negro and Schorfheide (2009) and Cúrdia and Reis (2012).

parameter, λ. If λ is equal to ∞, then one estimates an unrestricted VAR and if λ is equal to 0, then the procedure boils down to estimating a DSGE without allowing for misspecification. The estimation is executed for different values of λ. To determine the optimal value for λ, the authors propose using the marginal data density, which compares in-sample fit with model complexity.[88] If the restrictions imposed by the DSGE model are incorrect, then the procedure will put more weight on the VAR.

As pointed out in Chari et al. (2008), DSGE models often do not imply a VAR representation with a finite number of lags, unless all state variables are included. Thus, not only the DSGE, but also the VAR component of the DSGE-VAR procedure could be misspecified.

Wedges. Yet another approach to deal with misspecification is to add "wedges" to specific model equations. This procedure was introduced in Chari et al. (2007). Inoue et al. (2015) use this setup to formally test for model misspecfication. A wedge may have different interpretations or possibly no simple interpretation. From an econometric point a view, wedges are not different from regular structural disturbances in how they affect time series properties of the model. That is, they impose restrictions on the policy functions just as structural disturbances do and it matters crucially how one enters wedges. For example, the assumption that a wedge only enters one and not all model equations is a restriction. Although some wedges can enter more than one equation, wedges used in the literature only enter a few specific model equations and these are chosen by the researcher a priori and – as pointed out in Inoue et al. (2015) – wedges can be introduced in different ways. By contrast, ASDs appear in all equations and if one prefers a more concise specification, then our agnostic approach indicates one should use a statistical model selection criterion.

[88]The DSGE is less complex because it has fewer parameters, but could provide a worse in-sample fit, because of the restrictions it imposes.

3.8 Bibliography

An, S. and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26, 113–172.

Ascari, G., L. Phaneuf, and E. Sims (2016): "Business Cycles, Investment Shocks, and the "Barro-King Curse"," Department of Economics Discussion Paper Series, Number 815.

Beltran, D. O. and D. Draper (2015): "Estimating Dynamic Macroeconomic Models: How Informative are the Data?" Unpublished manuscript.

Canova, F., F. Ferroni, and C. Matthes (2015): "Approximating Time Varying Structural Models with Time Invariant Structures," Unpublished manuscript.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007): "Business Cycle Accounting," *Econometrica*, Forthcoming.

——— (2008): "Are Structural VARs with Long-Run Restrictions Useful in Developing Business Cycle Theory," *Journal of Monetary Economics*, 55.

Cúrdia, V. and R. Reis (2012): "Correlated Disturbances and U.S. Business Cycles," NBER Working Paper No. 15774.

Del Negro, M. and F. Schorfheide (2009): "Montetary Policy Analysis with Pontentially Misspecified Models," *American Economic Re*, 99, 1415–1450.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007): "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 25, 123–143.

Ferroni, F., S. Grassi, and M. A. León-Ledesma (2015): "Fundamental Shock Selection in DSGE models," Unpublished manuscript.

Gertler, M. and P. Karadi (2010): "A Model of Unconventional Monetary Policy," *Journal of Monetary Economics*, 58, 17–34.

Inoue, A., C.-H. Kuo, and B. Rossi (2015): "Identifying the Sources of Model Misspecification," Unpublished manuscript, Pompeu Fabra,.

Ireland, P. (2004): "Money's Role in the Monetary Business Cycle," *Journal of Money Credit and Banking*, 36, 969–983.

Iskrev, N. (2010): "Local Identification in DSGE Models," *Journal of Monetary Economics*, 57, 189–202.

Justiano, A., G. Primiceri, and A. Tambalotti (2010): "Investment Shocks and Business Cycles," *Journal of Monetary Economics*, 57, 132–145.

Komunjer, I. and S. Ng (2011): "Dynamic Identification of Dynamic Stochastic General Equilibrium Models," *Econometrica*, 79, 1995–2032.

Mickelsson, G. (2015): "Estimation of DSGE models: Maximum Likelihood vs. Bayesian Methods," Uppsala University Department of Economics Working Paper 2015:6.

Onatski, A. and N. Williams (2010): "Empirical and Policy Performance of a Forward-Looking Monetary Model," *Journal of Applied Econometrics*, 25, 145–176.

Paccagnini, A. (2017): "Dealing with Misspecificaiton in DSGE Models: A Survey," Munich Personal RePec Archive Paper No. 82914.

Smets, F. and R. Wouters (2003): "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123–1175.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

Chapter 4

Tracking the Slowdown in Long-Run GDP Growth

4.1 Introduction

> *"The global recovery has been disappointing (...) Year after year we have had to explain from mid-year on why the global growth rate has been lower than predicted as little as two quarters back".* Stanley Fischer, August 2014.

The slow pace of the recovery from the Great Recession of 2007-2009 has prompted questions about whether the long-run growth rate of GDP in advanced economies is lower now than it has been on average over the past decades (see e.g. Fernald, 2014, Gordon, 2014b, Summers, 2014). Indeed, forecasts of US and global real GDP growth have been persistently too optimistic for the last six years.[1] As emphasized by Orphanides (2003), real-time misperceptions about the long-run growth of the economy can play a large role in monetary policy mistakes. Moreover, small changes in assumptions about the long-run growth rate of output can have large implications on fiscal sustainability calculations (Auerbach, 2011). This calls for a framework that takes the uncertainty about long-run growth seriously and can inform decision-making in real time. In this paper, we present a dynamic factor model (DFM) which allows for gradual changes in the mean and the variance of real output growth. By incorporating a broad panel of economic activity indicators, DFMs are capable of precisely estimating the cyclical comovement in macroeconomic data in a real-time setting. Our model exploits this to track changes in the long-run growth rate of real GDP in a timely and reliable manner, separating them from their cyclical counterpart.[2]

The evidence of a decline in long-run US growth is accumulating, as documented

[1] For instance, Federal Open Market Committee (FOMC) projections since 2009 expected US growth to accelerate substantially, only to downgrade the forecast back to 2% throughout the course of the subsequent year. An analysis of forecasts produced by international organizations and private sector economists reveals the same pattern, see Pain et al. (2014) for a retrospective.

[2] Throughout this paper, our concept of the long run refers to changes in growth that are permanent in nature, i.e. do not mean-revert, as in Beveridge and Nelson (1981). In practice this should be thought of as frequencies lower than the business cycle.

by the recent growth literature such as Fernald and Jones (2014). Lawrence Summers and Robert Gordon have articulated a particularly pessimistic view of long-run growth which contrasts with the optimism prevailing before the Great Recession (see Jorgenson et al., 2006). To complement this evidence, we start our analysis by presenting the results of two popular structural break tests proposed by Nyblom (1989) and Bai and Perron (1998). Both suggest that a possible shift in the mean of US real GDP growth exists, the latter approach suggesting that a break probably occurred in the early part of the 2000's.[3] However, sequential testing using real-time data reveals that the break would not have been detected at conventional significance levels until as late as mid-2012, highlighting the problems of conventional break tests for real-time analysis (see also Benati, 2007). To address this issue, we introduce two novel features into an otherwise standard DFM of real activity data. First, we allow the mean of real GDP growth, and possibly other series, to drift gradually over time. As emphasized by Cogley (2005), if the long-run output growth rate is not constant, it is optimal to give more weight to recent data when estimating its current state. By taking a Bayesian approach, we can combine our prior beliefs about the rate at which the past information should be discounted with the information contained in the data. We also characterize the uncertainty around estimates of long-run growth taking into account both filtering and parameter uncertainty. Second, we allow for stochastic volatility (SV) in the innovations to both factors and idiosyncratic components. Given our interest in studying the entire postwar period, the inclusion of SV is essential to capture the substantial changes in the volatility of output that have taken place in this sample, such as the "Great Moderation" first reported by Kim and Nelson (1999a) and McConnell and Perez-Quiros (2000), as well as the cyclicality of macroeconomic volatility as documented by Jurado et al. (2014).

When applied to US data, our model concludes that long-run GDP growth declined meaningfully during the 2000's and currently stands at about 2%, more than one percentage point lower than the postwar average. The results are supportive of a gradual decline rather than a discrete break. Since in-sample results obtained with revised data often underestimate the uncertainty faced by policymakers in real time, we repeat the exercise using real-time vintages of data. The model detects the fall from the beginning of the 2000's onwards, and by the summer of 2010 it reaches the significant conclusion that a decline in long-run growth is behind the slow recovery, well before the structural break tests become conclusive.

We also investigate the performance of the model in "nowcasting" short-term developments in GDP. Since the seminal contributions of Evans (2005) and Giannone et al. (2008) DFMs have become the standard tool for this purpose.[4] Interestingly, our analysis shows that standard DFM forecasts revert very quickly to the unconditional mean of GDP, so taking into account the variation in long-run GDP growth

[3] This finding is consistent with the analysis of US real GDP by Luo and Startz (2014), as well as Fernald (2014), who applies the Bai and Perron (1998) test to US labor productivity.

[4] An extensive survey of the nowcasting literature is provided by Banbura et al. (2012), who also demonstrate, in a real-time context, the good out-of-sample performance of DFM nowcasts.

substantially improves point and density GDP forecasts even at very short horizons.

Finally, we extend our model in order to disentangle the drivers of secular fluctuations of GDP growth. Edge et al. (2007) emphasize the relevance as well as the difficulty of tracking permanent shifts in productivity growth in real time. In our framework, long-run output growth can be decomposed into labor productivity and labor input trends. The results of this decomposition exercise point to a slowdown in labor productivity as the main driver of recent weakness in GDP growth. Applying the model to other advanced economies, we provide evidence that the weakening in labor productivity appears to be a global phenomenon.

Our work is closely related to two strands of literature. The first one encompasses papers that allow for structural changes within the DFM framework. Del Negro and Otrok (2008) model time variation in factor loadings and volatilities, while Marcellino et al. (2014) show that the addition of SV improves the performance of the model for short-term forecasting of euro area GDP.[5] Acknowledging the importance of allowing for time-variation in the means of the variables, Stock and Watson (2012) pre-filter their data set in order to remove any low-frequency trends from the resulting growth rates using a biweight local mean. In his comment to their paper, Sims (2012) suggests to explicitly model, rather than filter out, these long-run trends, and emphasizes the importance of evolving volatilities for describing and understanding macroeconomic data. We see the present paper as extending the DFM literature, and in particular its application to tracking GDP, in the direction suggested by Chris Sims. The second strand of related literature takes a similar approach to decomposing long-run GDP growth into its drivers, in particular Gordon (2010, 2014a) and Reifschneider et al. (2013). Relative to these studies, we emphasize the importance of using a broader information set, as well as a Bayesian approach, which allows to use priors to inform the estimate of long-run growth, and to characterize the uncertainty around the estimate stemming both from filtering and parameter uncertainty.

The remainder of this paper is organized as follows. Section 4.2 presents preliminary evidence of a slowdown in long-run US GDP growth. Section 4.3 discusses the implications of time-varying long-run output growth and volatility for DFMs and presents our model. Section 4.4 applies the model to US data and documents the decline in long-run growth. The implications for tracking GDP in real time as well as the key advantages of our methodology are discussed. Section 4.5 decomposes the changes in long-run output growth into its underlying drivers. Section 4.6 concludes.

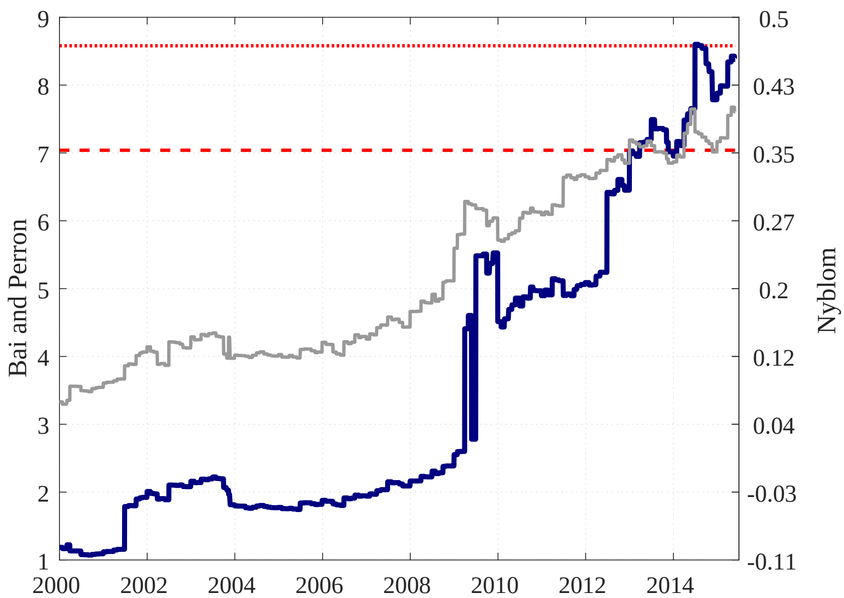
4.2 Preliminary Evidence

The literature on economic growth favors a view of the long-run growth rate as a process that evolves over time. It is by now widely accepted that a slowdown in productivity and long-run output growth occurred in the early 1970's, and that ac-

[5] While the model of Del Negro and Otrok (2008) includes time-varying factor loadings, the means of the observable variables are still treated as constant.

celerating productivity in the IT sector led to a boom in the late 1990's.[6] In contrast, in the context of econometric modeling the possibility that long-run growth is time-varying is the source of a long-standing controversy. In their seminal contribution, Nelson and Plosser (1982) model the (log) level of real GDP as a random walk with drift. This implies that after first-differencing, the resulting growth rate fluctuates around a constant mean, an assumption still embedded in many econometric models. After the slowdown in productivity became apparent in the 1970's, many researchers such as Clark (1987) modeled the drift term as an additional random walk, implying that the level of GDP is integrated of order two. The latter assumption would also be consistent with the local linear trend model of Harvey (1985), the Hodrick and Prescott (1997) filter, and Stock and Watson (2012)'s practice of removing a local biweight mean from the growth rates before estimating a DFM. The $I(2)$ assumption is nevertheless controversial since it implies that the growth rate of output can drift without bound. Consequently, papers such as Perron and Wada (2009), have modeled the growth rate of GDP as stationary around a trend with one large break around 1973.

Figure 4.1: Real-Time Test Statistics of Nyblom and Bai-Perron Tests



Note: The gray and blue solid lines are the values of the test statistics obtained from sequentially re-applying the Nyblom (1989) and Bai and Perron (1998) tests in real time as new National Accounts vintages are being published. In both cases, the sample starts in 1947:Q2 and the test is re-applied for every new data release occurring after the beginning of 2000. The dotted and dashed red lines represent the 5% and 10% critical values corresponding to the two tests.

Ever since the Great Recession of 2007-2009 US real GDP has grown well below its postwar average, once again raising the question whether its mean may have declined. There are two popular strategies that could be followed from a frequentist perspective to detect parameter instability or the presence of breaks in the mean

[6]For a retrospective on the productivity slowdown, see Nordhaus (2004). Oliner and Sichel (2000) provide evidence on the role of the IT sector in the acceleration of the late 1990's.

growth rate. The first one is Nyblom's (1989) L-test as described in Hansen (1992), which tests the null hypothesis of constant parameters against the alternative that the parameters follow a martingale. Modeling real GDP growth as an AR(1) over the sample 1947-2015 this test rejects the stability of the constant term at the 10% significance level.[7] The second commonly used approach, which can determine the number and timing of multiple discrete breaks, is the Bai and Perron (1998) test. This test finds evidence in favor of a single break in the mean of US real GDP growth at the 10%-level. The most likely break date is in the second quarter of 2000. In related research, Fernald (2014) provides evidence for breaks in labor productivity in 1973:Q2, 1995:Q3, and 2003:Q1, and links the latter two to developments in the IT sector. From a Bayesian perspective, Luo and Startz (2014) calculate the posterior probability of a single break and find the most likely break date to be 2006:Q1 for the full postwar sample and 1973:Q1 for a sample excluding the 2000's.

The above results indicate that substantial evidence for a recent change in the mean of US GDP growth has built up. However, the strategy of applying conventional tests and introducing deterministic breaks into econometric models is not satisfactory for the purposes of real-time decision making. In fact, the detection of change in the mean of GDP growth can arrive with substantial delay. To demonstrate this, a sequential application of the Nyblom (1989) and Bai and Perron (1998) tests using real-time data is presented in Figure 4.1. The evolution of the test statistics in real-time reveals that a break would not have been detected at the 10% significance levels until as late as mid-2012, which is more than ten years later than the actual break date suggested by the Bai and Perron (1998) procedure. The Nyblom (1989) test, which is designed to detect gradual change rather than a discrete break, becomes significant roughly at the same time. This lack of timeliness highlights the importance of an econometric framework capable of quickly adapting to changes in long-run growth as new information arrives.

4.3 Econometric Framework

DFMs in the spirit of Geweke (1977), Stock and Watson (2002) and Forni et al. (2009) capture the idea that a small number of unobserved factors drives the comovement of a possibly large number of macroeconomic time series, each of which may be contaminated by measurement error or other sources of idiosyncratic variation. Their theoretical appeal (see e.g. Sargent and Sims, 1977 or Giannone et al., 2006), as well as their ability to parsimoniously model large data sets, have made them a workhorse of empirical macroeconomics. Giannone et al. (2008) and Banbura et al. (2012) have pioneered the use of DFMs to produce current-quarter forecasts ("nowcasts") of GDP growth by exploiting more timely monthly indicators and the factor structure of the data. Given the widespread use of DFMs to track GDP in real time, this paper aims

[7]The same result holds for an AR(2) specification. In both cases, stability of the autoregressive coefficients cannot be rejected, whereas stability of the variance is rejected at the 1%-level. Section 4.7.2 of the Appendix provides the full results of both tests applied in this section.

to make these models robust to changes in long-run growth. We do so by introducing two novel features into the DFM framework. First, we allow the long-run growth rate of real GDP, and possibly other series, to vary over time. Second, we allow for stochastic volatility (SV) in the innovations to both factors and idiosyncratic components, given our interest in studying the entire postwar period for which drastic changes in volatility have been documented. With these changes, the DFM proves to be a powerful tool to detect changes in long-run growth. The information contained in a broad panel of activity indicators facilitates the timely decomposition of real GDP growth into persistent long-run movements, cyclical fluctuations and short-lived noise.

4.3.1 The Model

Let $\mathbf{y}_t$ be an $n \times 1$ vector of observable macroeconomic time series, and let $\mathbf{f}_t$ denote a $k \times 1$ vector of latent common factors. It is assumed that $n >> k$, i.e. the number of observables is much larger than the number of factors. Formally,

$$\mathbf{y}_t = \mathbf{c}_t + \boldsymbol{\Lambda}\mathbf{f}_t + \mathbf{u}_t, \tag{4.1}$$

where $\boldsymbol{\Lambda}$ contains the loadings on the common factors and $\mathbf{u}_t$ is a vector of idiosyncratic components.[8] Shifts in the long-run mean of $\mathbf{y}_t$ are captured by time-variation in $\mathbf{c}_t$. In principle one could allow time-varying intercepts in all or a subset of the variables in the system. Moreover, time variation in a given series could be shared by other series. $\mathbf{c}_t$ is therefore flexibly specified as

$$\mathbf{c}_t = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{c} \end{bmatrix} \begin{bmatrix} \mathbf{a}_t \\ 1 \end{bmatrix}, \tag{4.2}$$

where $\mathbf{a}_t$ is an $r \times 1$ vector of time-varying means, $\mathbf{B}$ is an $m \times r$ matrix which governs how the time-variation affects the corresponding observables, and $\mathbf{c}$ is an $(n-m) \times 1$ vector of constants. In our baseline specification, $\mathbf{a}_t$ will be a scalar capturing time-variation in long-run real GDP growth, which is shared by real consumption growth, so that $r = 1, m = 2$. A detailed discussion of this and additional specifications of $\mathbf{c}_t$ will be provided in Section 4.3.2. Throughout the paper, we focus on the case of a single dynamic factor by setting $k = 1$ (i.e. $\mathbf{f}_t = f_t$).[9] The laws of motion of the

[8]The model can be extended to include lags of the factor in the measurement equation. In the latter case, it is sensible to avoid overfitting by choosing priors that shrink the additional lag coefficients towards zero (see D'Agostino et al., 2015, and Luciani and Ricci, 2014). We consider this possibility when we explore robustness of our results to using larger data panels in Section 4.4.6.

[9]For the purpose of tracking real GDP with a large number of closely related activity indicators, the use of one factor is appropriate, which is explained in more detail in Sections 4.4.1 and 4.4.2. Also note that we order real GDP growth as the first element of $\mathbf{y}_t$, and normalize the loading for GDP to unity. This serves as an identifying restriction in our estimation algorithm. Bai and Wang (2015) discuss minimal identifying assumptions for DFMs.

latent factor and the idiosyncratic components are

$$(1-\phi(L))f_t = \sigma_{\varepsilon_t}\varepsilon_t, \tag{4.3}$$

$$(1-\rho_i(L))u_{i,t} = \sigma_{\eta_{i,t}}\eta_{i,t}, \quad i=1,\dots,n \tag{4.4}$$

where $\phi(L)$ and $\rho_i(L)$ denote polynomials in the lag operator of order p and q, respectively. The idiosyncratic components are cross-sectionally orthogonal and are assumed to be uncorrelated with the common factor at all leads and lags, i.e. $\varepsilon_t \overset{iid}{\sim} N(0,1)$ and $\eta_{i,t} \overset{iid}{\sim} N(0,1)$.

Finally, the dynamics of the model's time-varying parameters are specified to follow driftless random walks:

$$a_{j,t} = a_{j,t-1} + v_{a_{j,t}}, \qquad v_{a_{j,t}} \overset{iid}{\sim} N(0,\omega_{a,j}^2) \quad j=1,\dots,r \tag{4.5}$$

$$\log\sigma_{\varepsilon_t} = \log\sigma_{\varepsilon_{t-1}} + v_{\varepsilon,t}, \qquad v_{\varepsilon,t} \overset{iid}{\sim} N(0,\omega_{\varepsilon}^2) \tag{4.6}$$

$$\log\sigma_{\eta_{i,t}} = \log\sigma_{\eta_{i,t-1}} + v_{\eta_{i,t}}, \qquad v_{\eta_{i,t}} \overset{iid}{\sim} N(0,\omega_{\eta,i}^2) \quad i=1,\dots,n \tag{4.7}$$

where $a_{j,t}$ are the r time-varying elements in $\mathbf{a}_t$, and σ_{ε_t} and $\sigma_{\eta_{i,t}}$ capture the SV of the innovations to factor and idiosyncratic components. Our motivation for specifying the time-varying parameters as random walks is similar to Primiceri (2005). While in principle it is unrealistic model real GDP growth as a process that could wander in an unbounded way, as long as the variance of the process is small and the drift is considered to be operating for a finite period of time, the assumption is innocuous. Moreover, modeling a trend as a random walk is more robust to misspecification when the actual process is instead characterized by discrete breaks, whereas models with discrete breaks might not be robust to the true process being a random walk.[10] Finally, the random walk assumption also has the desirable feature that, unlike stationary models, confidence bands around forecasts of real GDP growth increase with the forecast horizon, reflecting uncertainty about the possibility of future breaks or drifts in long-run growth.

Note that a standard DFM is usually specified under two assumptions. First, the original data have been differenced appropriately so that both the factor and the idiosyncratic components can be assumed to be stationary. Second, it is assumed that the innovations in the idiosyncratic and common components are iid. In equations (4.1)-(4.7) we have relaxed these assumptions to allow for two novel features, a stochastic trend in the mean of selected series, and SV. By shutting down these features, we can recover the specifications previously proposed in the literature, which are nested in our framework. We obtain the DFM with SV of Marcellino et al. (2014) if we shut down time-variation in the intercepts of the observables, i.e. set $r=m=0$ and $\mathbf{c}_t = \mathbf{c}$. If we further shut down the SV, i.e. set $\omega_{a,j}^2 = \omega_{\epsilon}^2 = \omega_{\eta,i}^2 = 0$, we obtain

[10] We demonstrate this point with the use of Monte Carlo simulations, showing that a random walk trend in real GDP growth 'learns' quickly about a discrete break once it has occurred. On the other hand, the random walk does not detect a drift when there is not one, despite the presence of a large cyclical component. Appendix 4.7.3 provides a discussion and the full results of these simulations.

the specification of Banbura and Modugno (2014) and Banbura et al. (2012).

4.3.2 A Baseline Specification for Long-Run Growth

Equations (4.1) and (4.2) allow for stochastic trends in the mean of all or a subset of selected observables in $\mathbf{y}_t$. This paper focuses on tracking changes in the long-run growth rate of real GDP. For this purpose, the simplest specification of $\mathbf{c}_t$ is to include a time-varying intercept only in GDP and to set $\mathbf{B} = 1$. However, a number of empirical studies (e.g. Harvey and Stock, 1988, Cochrane, 1994, and Cogley, 2005) argue that incorporating information about consumption is informative about the permanent component in GDP as predicted by the permanent income hypothesis. The theory predicts that consumers, smoothing consumption throughout their lifetime, should react more strongly to permanent, as opposed to transitory, changes in income. As a consequence, looking at GDP and consumption data together will help separating growth into long-run and cyclical fluctuations.[11] Therefore, our baseline specification imposes that consumption and output grow at the same rate g_t in the long-run. On the contrary, we do not impose that investment also grows at this rate, as would be the case in the basic neoclassical growth model, since the presence of investment-specific technological change implies that real investment has a different low-frequency trend (Greenwood et al., 1997).

Formally, ordering real GDP and consumption growth first, and setting $m = 2$ and $r = 1$, this is represented as

$$\mathbf{a}_t = g_t, \quad \mathbf{B} = [1\ 1]' \tag{4.8}$$

Note that in this baseline specification we model time-variation only in the intercept for GDP and consumption while leaving it constant for the other observables. Of course it may be the case that some of the remaining $n - m$ series in $\mathbf{y}_t$ feature low frequency variation in their means. For instance, as mentioned above, this could be the case for investment. The key question is whether leaving it unspecified will affect the estimate of the long-run growth rate of GDP, which is our main object of interest. We ensure that this is not the case by allowing for persistence (and, in particular, we do not rule out unit roots) in the idiosyncratic components. If a series does feature a unit root which is not included in $\mathbf{a}_t$, its trend component will be absorbed by the idiosyncratic component. The choice of which elements to include in $\mathbf{a}_t$ therefore reflects the focus of a particular application.[12] Of course, if

[11] While a strict interpretation of the permanent income hypothesis is rejected in the data, from an econometric point of view the statement applies as long as permanent changes are the main driver of consumption. See Cochrane (1994) for a very similar discussion.

[12] In principle, these unmodeled trends could still be recovered from our specification by applying a Beveridge-Nelson decomposition to its estimated idiosyncratic component. In practice, any low-frequency variation in the idiosyncratic component is likely to be obscured by a large amount of high frequency noise in the data and as result the extracted Beveridge-Nelson trend component will be imprecisely estimated, and as Morley et al. (2003) show, will not be smooth. In our specification, the elements of $\mathbf{a}_t$ are instead extracted directly, so that we are able to improve the extraction by imposing additional assumptions (e.g. smoothness) and prior beliefs (e.g. low variability) on its properties.

two series share the same underlying low-frequency component, and this is known with certainty, explicitly accounting for the shared low frequency variation will improve the precision of the estimation, but the risk of incorrectly including the trend is much larger than the risk of incorrectly excluding it. Therefore, in our baseline specification we include in $\mathbf{a}_t$ the intercept for GDP and consumption, while leaving any possible low-frequency variation in other series to be captured by the respective idiosyncratic components.[13]

An extension to include additional time-varying intercepts is straightforward through the flexible construction of $\mathbf{c}_t$ in equation (4.2). In fact, in Section 4.5 we explore how interest in the low-frequency movements of additional series leads to alternative choices for $\mathbf{a}_t$ and $\mathbf{B}$.[14]

4.3.3 Dealing with Mixed Frequencies and Missing Data

Tracking activity in real time requires a model that can efficiently incorporate information from series measured at different frequencies. In particular, it must include both quarterly variables, such as the growth rate of real GDP, as well as more timely monthly indicators of real activity. Therefore, the model is specified at monthly frequency, and following Mariano and Murasawa (2003), the (observed) quarterly growth rates of a generic quarterly variable, x_t^q, can be related to the (unobserved) monthly growth rate x_t^m and its lags using a weighted mean. Specifically,

$$x_t^q = \frac{1}{3}x_t^m + \frac{2}{3}x_{t-1}^m + x_{t-2}^m + \frac{2}{3}x_{t-3}^m + \frac{1}{3}x_{t-4}^m, \tag{4.9}$$

and only every third observation of x_t^q is actually observed. Substituting the corresponding line of (4.1) into (4.9) yields a representation in which the quarterly variable depends on the factor and its lags. The presence of mixed frequencies is thus reduced to a problem of missing data in a monthly model.

Besides mixed frequencies, additional sources of missing data in the panel include: the "ragged edge" at the end of the sample, which stems from the non-synchronicity of data releases; missing data at the beginning of the sample, since some data series have been created or collected more recently than others; and missing observations due to outliers and data collection errors. Our Bayesian estimation method exploits the state space representation of the DFM and jointly estimates the latent factors, the parameters, and the missing data points using the Kalman filter (see Durbin and Koopman, 2012, for a textbook treatment).

[13] We confirm this line of reasoning with a series of Monte Carlo experiments, in which data is generated from a system that features low-frequency movements in more series, which are left unmodeled in the estimation. Both in the case of series with independent trends and the case of series which share the trend of interest, the fact that they are left unmodeled has little impact on the estimate of the latter. Appendix 4.7.3 presents further discussion and the full results of these simulations.

[14] The limiting case explicitly models time-varying intercept in all indicators, so that $m = r = n$ and $\mathbf{B} = \mathbf{I}_n$, i.e. an identity matrix of dimension n. See Creal et al. (2010) and Fleischman and Roberts (2011) for similar approaches. This setup would imply that the number of state variables increases with the number of observables, which severely increases the computational burden of the estimation, while offering little additional evidence with respect to the focus of this paper.

4.3.4 State Space Representation and Estimation

The model features autocorrelated idiosyncratic components (see equation (4.4)). In order to cast it in state-space form, we include the idiosyncratic components of the quarterly variables in the state vector, and we redefine the system for the monthly indicators in terms of quasi-differences (see e.g. Kim and Nelson, 1999b, pp. 198-199, and Bai and Wang, 2015).[15] The model is estimated with Bayesian methods simulating the posterior distribution of parameters and factors using a Markov Chain Monte Carlo (MCMC) algorithm. We closely follow the Gibbs-sampling algorithm for DFMs proposed by Bai and Wang (2015), but extend it to include mixed frequencies, the time-varying intercept, and SV. The SVs are sampled using the approximation of Kim et al. (1998), which is considerably faster than the exact Metropolis-Hastings algorithm of Jacquier et al. (2002). Our complete sampling algorithm together with the details of the state space representation can be found in Section 4.7.4 of the Appendix.

4.4 Results for US Data

4.4.1 Data Selection

Our data set includes four key business cycle variables measured at quarterly frequency (output, consumption, investment and aggregate hours worked), as well as a set of 24 monthly indicators which are intended to provide additional information about cyclical developments in a timely manner.

The included quarterly variables are strongly procyclical and are considered key indicators of the business cycle (see e.g. Stock and Watson, 1999). Furthermore, theory predicts that they will be useful in disentangling low frequency movements from cyclical fluctuations in output growth. Indeed, as discussed in Section 4.3.2, the permanent income hypothesis predicts that consumption data will be particularly useful for the estimation of the long-run growth component, g_t.[16] On the other hand, investment and hours worked are very sensitive to cyclical fluctuations, and thus will be particularly informative for the estimation of the common factor, f_t.[17]

[15] Since the quarterly variables are observed only every third month, we cannot take the quasi-difference for their idiosyncratic components, which are instead added as an additional state with the corresponding transition dynamics. Banbura and Modugno (2014) suggest including all of the idiosyncratic components as additional elements of the state vector. Our solution has the desirable feature that the number of state variables will increase with the number of quarterly variables, rather than the total number of variables, leading to a gain of computational efficiency.

[16] Due to the presence of faster technological change in the durable goods sector there is a downward trend in the relative price of durable goods. As a consequence, measured consumption grows faster than overall GDP. Following a long tradition in the literature (see e.g. Whelan, 2003), we construct a Fisher index of non-durables and services and use its growth rate as an observable variable in the panel. It can be verified that the ratio of consumption defined in this manner to real GDP displays no trend in the data, unlike the trend observed in the ratio of overall consumption to GDP.

[17] We define investment as a chain-linked aggregate of business fixed investment and consumption of durable goods, which is consistent with our treatment of consumption. In order to obtain a measure of hours for the total economy, we follow the methodology of Ohanian and Raffo (2012)

The additional monthly indicators are crucial to our objective of disentangling in real time the cyclical and long-run components of GDP growth, since the quarterly variables are only available with substantial delay. In principle, a large number of candidate series are available to inform the estimate of f_t, and indirectly, of g_t. In practice, however, macroeconomic data series are typically clustered in a small number of broad categories (such as production, employment, or income) for which disaggregated series are available along various dimensions (such as economic sectors, demographic characteristics, or expenditure categories). The choice of which available series to include for estimation can therefore be broken into, first, a choice of which broad categories to include, and second, to which level and along which dimensions of disaggregation.

With regards to which broad categories of data to include, previous studies agree that prices, monetary and financial indicators are uninformative for the purpose of tracking real GDP, and argue for extracting a single common factor that captures real economic activity.[18] As for the possible inclusion of disaggregated series within each category, Boivin and Ng (2006) argue that the presence of strong correlation in the idiosyncratic components of disaggregated series of the same category will be a source of misspecification that can worsen the performance of the model in terms of in-sample fit and out-of-sample forecasting of key series.[19] Alvarez et al. (2012) investigate the trade-off between DFMs with very few indicators, where the good large-sample properties of factor models are unlikely to hold, and those with a very large amount of indicators, where the problems above are likely to arise. They conclude that using a medium-sized panel with representative indicators of each category yields the best forecasting results.

The above considerations lead us to select 24 monthly indicators that include the high-level aggregates for all of the available broad categories that capture real activity, without overweighting any particular category. The complete list of variables contained in our data set is presented in Table 4.1. As the table shows, we include representative series of expenditure and income, the labor market, production and sales, foreign trade, housing and business and consumer confidence.[20] The inclusion of all the available monthly surveys is particularly important. Apart from being the most timely series available, these are unlikely to feature permanent shifts in their mean by construction, and have a high signal-to-noise ratio. They thus provide a

and benchmark the quarterly series of hours in the non-farm business sector provided by the BLS to the annual estimates of hours in the total economy compiled by the Conference Board's Total Economy Database (TED). The TED series has the advantage of being comparable across countries (Ohanian and Raffo, 2012), which will be useful for our international results in Section 4.5.

[18] Giannone et al. (2005) conclude that that prices and monetary indicators do not contribute to the precision of GDP nowcasts. Banbura et al. (2012), Forni et al. (2003) and Stock and Watson (2003) find at best mixed results for financial variables.

[19] This problem is exacerbated by the fact that more detailed disaggregation levels and dimensions are available for certain categories of data, such as employment, meaning that the disaggregation will automatically 'tilt' the factor estimates towards that category.

[20] When there are multiple candidates for the high-level aggregate of a category, we include both. For example, we include employment as measured both by the establishment and household surveys, and consumer confidence as surveyed both by the Conference Board and the University of Michigan.

clean signal to separate the cyclical component of GDP growth from its long-run counterpart. In Section 4.4.6 we explore sensitivity of our results to the size and composition of the data panel used.

Our panel spans the period January 1947 to March 2015. The start of our sample coincides with the year for which quarterly national accounts data are available from the Bureau of Economic Analysis. This enables us to study the evolution of long-run growth over the entire postwar period.[21]

4.4.2 Model Settings and Priors

The choice of the data set justifies the single-factor structure of the model. f_t can in this case be interpreted as a coincident indicator of real economic activity (see e.g. Stock and Watson, 1989, and Mariano and Murasawa, 2003). The number of lags in the polynomials $\phi(L)$ and $\rho(L)$ is set to $p = 2$ and $q = 2$ as in Stock and Watson (1989). We wish to impose as little prior information as possible, so we use uninformative priors for the factor loadings and the autoregressive coefficients of factors and idiosyncratic components. The variances of the innovations to the time-varying parameters, namely ω_a^2, ω_ε^2 and $\omega_{\eta,i}^2$ in equations (4.5)-(4.7) are however difficult to identify from the information contained in the likelihood alone. As the literature on Bayesian VARs documents, attempts to use non-informative priors for these parameters will in many cases produce posterior estimates which imply a relatively large amount of time-variation. While this will tend to improve the in-sample fit of the model it is also likely to worsen out-of-sample forecast performance. We therefore use priors to shrink these variances towards zero, i.e. towards the standard DFM which excludes time-varying long-run GDP growth and SV. In particular, for ω_a^2 we set an inverse gamma prior with one degree of freedom and scale equal to 0.001.[22] For ω_ϵ^2 and $\omega_{\eta,i}^2$ we set an inverse gamma prior with one degree of freedom and scale equal to 0.0001, closely following Cogley and Sargent (2005) and Primiceri (2005).[23] We estimate the model with 7000 replications of the Gibbs-sampling algorithm, of which the first 2000 are discarded as burn-in draws and the remaining ones are kept for inference.[24]

[21] We take full advantage of the Kalman filter's ability to deal with missing observations at any point in the sample, and we are able to incorporate series that become available substantially later than 1947, up to as late as 2007. Note that for consumption expenditures, monthly data became available in 1959, whereas quarterly data is available from 1947. In order to use all available data, we apply the polynomial in Equation (4.9) to the monthly data and treat the series as quarterly, with available observations for the last month of the quarter for 1947-1958 and for all months since 1959.

[22] To gain an intuition about this prior, note that over a period of ten years, this would imply that the random walk process of the long-run growth rate is expected to vary with a standard deviation of around 0.4 percentage points in annualized terms, which is a fairly conservative prior.

[23] We provide further explanations and address robustness to the choice of priors in Appendix 4.7.7.

[24] Thanks to the efficient state space representation discussed above, the improvements in the simulation smoother proposed by Bai and Wang (2015), and other computational improvements we implemented, the estimation is very fast. Convergence is achieved after only 1500 iterations, which take less than 20 minutes in MATLAB using an Intel 3.6 GHz computer with 16GB of DDR3 Ram.

Table 4.1: Data series used in empirical analysis

	Type	Start Date	Transform.	Lag
Quarterly time series				
Real GDP	Expenditure & Inc.	Q2:1947	% QoQ Ann	26
Real Consumption (excl. durables)	Expenditure & Inc.	Q2:1947	% QoQ Ann	26
Real Investment (incl. durable cons.)	Expenditure & Inc.	Q2:1947	% QoQ Ann	26
Total Hours Worked	Labor Market	Q2:1948	% QoQ Ann	28
Monthly indicators				
Real Personal Income less Transfers	Expenditure & Inc.	Feb 59	% MoM	27
Industrial Production	Production & Sales	Jan 47	% MoM	15
New Orders of Capital Goods	Production & Sales	Mar 68	% MoM	25
Real Retail Sales & Food Services	Production & Sales	Feb 47	% MoM	15
Light Weight Vehicle Sales	Production & Sales	Feb 67	% MoM	1
Real Exports of Goods	Foreign Trade	Feb 68	% MoM	35
Real Imports of Goods	Foreign Trade	Feb 69	% MoM	35
Building Permits	Housing	Feb 60	% MoM	19
Housing Starts	Housing	Feb 59	% MoM	26
New Home Sales	Housing	Feb 63	% MoM	26
Payroll Empl. (Establishment Survey)	Labor Market	Jan 47	% MoM	5
Civilian Empl. (Household Survey)	Labor Market	Feb 48	% MoM	5
Unemployed	Labor Market	Feb 48	% MoM	5
Initial Claims for Unempl. Insurance	Labor Market	Feb 48	% MoM	4
Monthly indicators (soft)				
Markit Manufacturing PMI	Business Confidence	May 07	-	-7
ISM Manufacturing PMI	Business Confidence	Jan 48	-	1
ISM Non-manufacturing PMI	Business Confidence	Jul 97	-	3
NFIB Small Business Optimism Index	Business Confidence	Oct 75	Diff 12 M.	15
U. of Michigan: Consumer Sentiment	Consumer Confid.	May 60	Diff 12 M.	-15
Conf. Board: Consumer Confidence	Consumer Confid.	Feb 68	Diff 12 M.	-5
Empire State Manufacturing Survey	Business (Regional)	Jul 01	-	-15
Richmond Fed Mfg Survey	Business (Regional)	Nov 93	-	-5
Chicago PMI	Business (Regional)	Feb 67	-	0
Philadelphia Fed Business Outlook	Business (Regional)	May 68	-	0

Note: % QoQ Ann refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. The last column shows the average publication lag, i.e. the number of days elapsed from the end of the period that the data point refers to until its publication by the statistical agency. All series were obtained from the Haver Analytics database.

4.4.3 In-Sample Results

Panel (a) of Figure 4.2 plots the posterior median, together with the 68% and 90% posterior credible intervals of the long-run growth rate of real GDP. This estimate is conditional on the entire sample and accounts for both filtering and parameter uncertainty. Several features of our estimate of long-run growth are worth noting. While the growth rate is stable between 3% and 4% during the first decades of the postwar period, a slowdown is clearly visible from around the late 1960's through the 1970's, consistent with the "productivity slowdown" (Nordhaus, 2004). The acceleration of the late 1990's and early 2000's associated with the productivity boom in the IT sector (Oliner and Sichel, 2000) is also visible. Thus, until the middle of the decade of the 2000's, our estimate conforms well to the generally accepted narrative about fluctuations in potential growth.[25] More recently, after peaking at about 3.5% in 2000, the median estimate of the long-run growth rate has fallen to about 2% in early 2015, a more substantial decline than the one observed after the productivity slowdown of the 1970's. Moreover, the slowdown appears to have happened gradually since the start of the 2000's, with most of the decline having occurred before the Great Recession.[26] Interestingly, a small rebound is visible at the end of the sample, but long-run growth stands far below its postwar average of 3.2%, with the 90% posterior credible interval ranging from 1.5% to 2.5%.

Panel (b) plots the time series of quarterly real GDP growth, together with the median posterior estimates of the common factor, aligned with the mean of real GDP growth. This plot highlights how the common factor captures the bulk of business-cycle frequency variation in output growth, while higher frequency, quarter-to-quarter variation is attributed to the idiosyncratic component. In the latter part of the sample, GDP growth is visibly below the factor, reflecting the decline in long-run growth.

The posterior estimate of the SV of the common factor is presented in Panel (c). It is clearly visible that volatility declines over the sample. The late 1940's and 1950's were extremely volatile, with a first large drop in volatility in the early 1960's. The Great Moderation is also clearly visible, with the average volatility pre-1985 being much larger than the average of the post-1985 sample. Notwithstanding the large increase in volatility during the Great Recession, our estimate of the common factor volatility since then remains consistent with the Great Moderation still being in place. This confirms the early evidence reported by Gadea-Rivas et al. (2014). It is clear from the figure that volatility spikes during recessions, a feature that brings

[25] Appendix 4.7.8 provides a comparison of our estimate with the Congressional Budget Office (CBO) measure of potential growth, with some additional discussion.

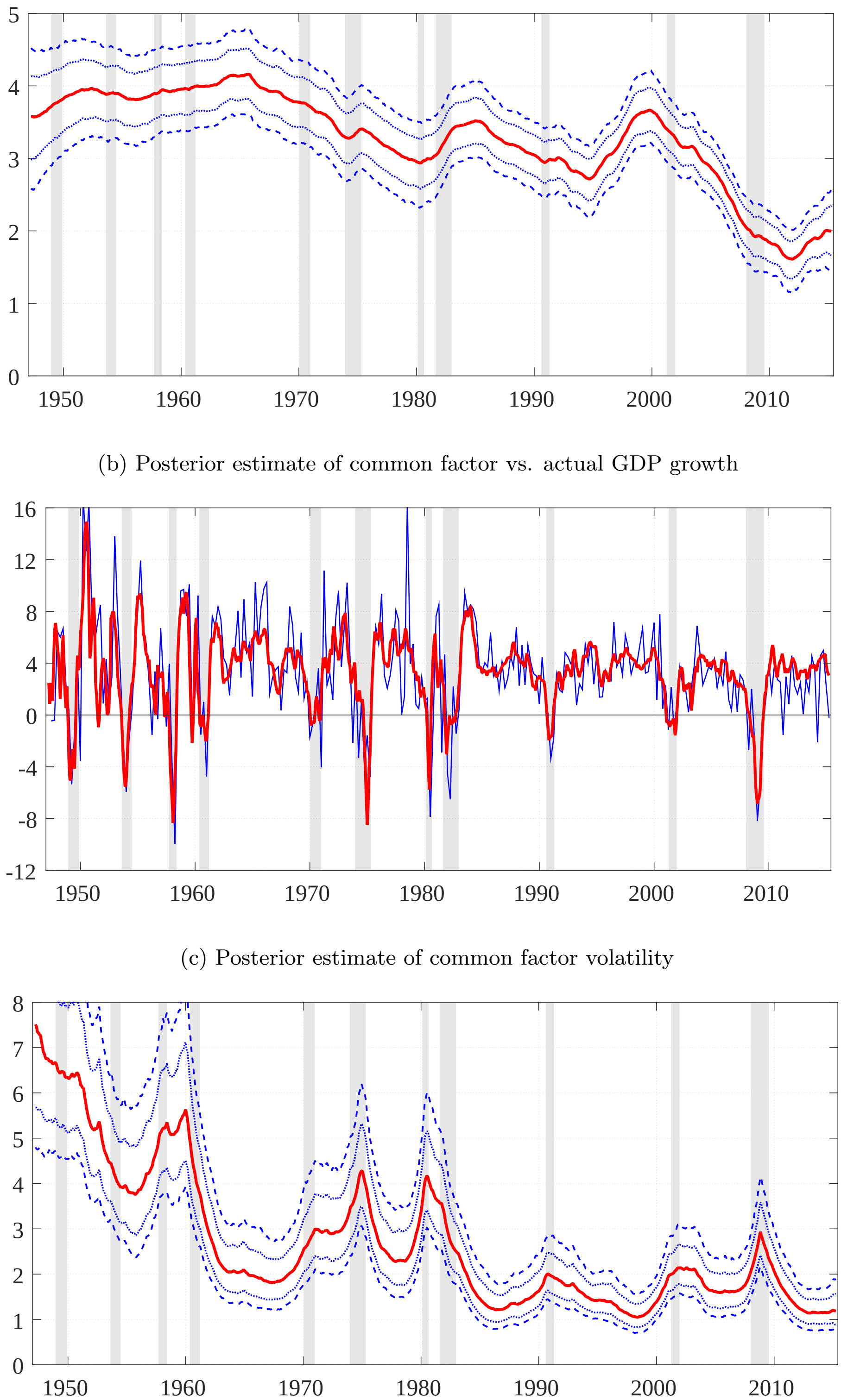
[26] In principle, it is possible that our choice of modeling long-run GDP growth as a random walk is hard-wiring into our results the conclusion that the decline happened in a gradual way. In experiments with simulated data, presented in Section 4.7.3 of the Appendix, we show that if changes in long-run growth occur in the form of discrete breaks rather than evolving gradually, the (one-sided) filtered estimates will exhibit a discrete jump at the moment of the break. Instead, for US data the filtered estimates of the long-run growth component also decline in a gradual manner (see Figure 4.6 in Appendix 4.7.1).

Figure 4.2: TREND, CYCLE AND VOLATILITY: 1947-2015 (% ANN.)

(a) Posterior estimate of long-run growth

(b) Posterior estimate of common factor vs. actual GDP growth

(c) Posterior estimate of common factor volatility



Note: Panel (a) displays the posterior median (solid red), together with the 68% and 90% (dotted and dashed blue) posterior credible intervals of long-run real GDP growth. Panel (b) plots actual real GDP growth (thin blue) against the posterior median estimate of the common factor, aligned with the mean of real GDP growth (thick red). Panel (c) presents the median (red), the 68% and the 90% (dotted and dashed blue) posterior credible intervals of the volatility of the common factor, i.e the square root of $\text{var}(f_t) = \sigma^2_{\varepsilon,t}(1-\phi_2)/[(1+\phi_2)((1-\phi_2)^2-\phi_1^2)]$. Shaded areas represent NBER recessions.

our estimates close to the recent findings of Jurado et al. (2014) and Bloom (2014) relating to business-cycle uncertainty.[27] It appears that the random walk specification is flexible enough to capture cyclical changes in volatility as well as permanent phenomena such as the Great Moderation. Appendix 4.7.1 contains analogous charts for the volatilities of the idiosyncratic components of selected data series. Similar to the volatility of the common factor, many of the idiosyncratic volatilities present sharp increases during recessions.

The above results provide evidence that a significant decline in long-run US real GDP growth occurred over the last decade, and are consistent with a relatively gradual decline since the early 2000's. Our estimates show that the bulk of the slowdown from the elevated levels of growth at the turn of the century occurred before the Great Recession, which is consistent with the narrative of Fernald (2014) on the fading of the IT productivity boom. This recent decline is the largest movement in long-run growth observed in the postwar period.

4.4.4 Real-Time Results

As emphasized by Orphanides (2003), macroeconomic time series are heavily revised over time and in many cases these revisions contain valuable information that was not available at initial release. Therefore, it is important to assess, using the data available at each point in time, when the model detected the slowdown in long-run growth. For this purpose, we reconstruct our data set using vintages of data available from the Federal Reserve Bank of St. Louis ALFRED data base. Our aim is to replicate as closely as possible the situation of a decision-maker which would have applied our model in real time. We fix the start of our sample in 1947:Q1 and use an expanding out-of-sample window which starts on 11 January 2000 and ends on 30 June 2015. This is the longest possible window for which we are able to include the entire panel in Table 4.1 using fully real-time data. We then proceed by re-estimating the model each day in which new data are released.[28]

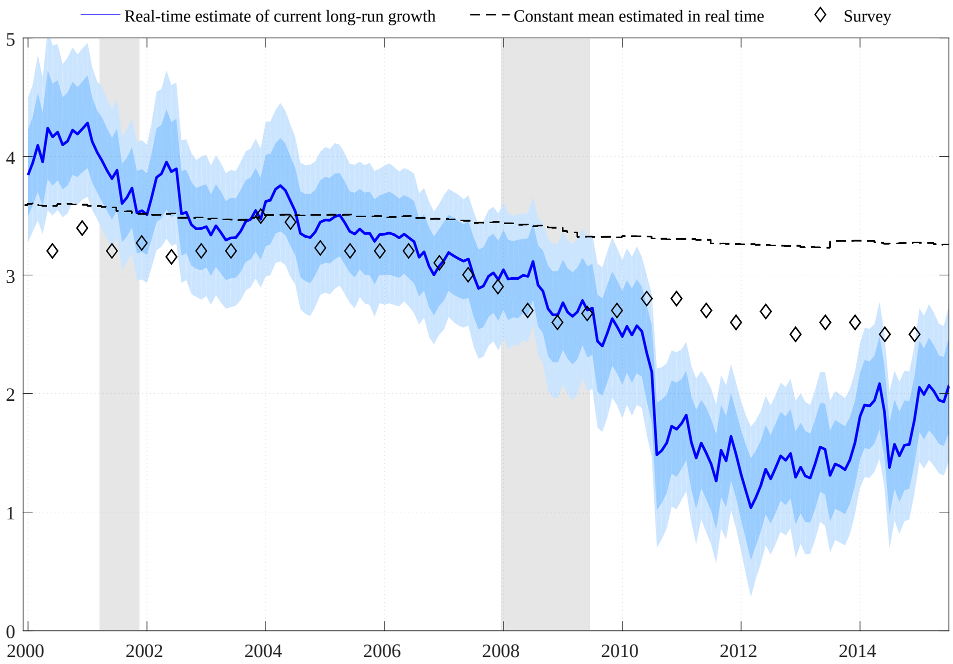
Figure 4.3 looks at the model's real-time assessment of long-run growth at various points in time. Panel (a) plots the real-time estimate of current long-run growth, with 68% and 90% uncertainty bands. For comparison, the panel also shows the median response to the Philadelphia Fed Livingston Survey of Professional Forecasters (SPF) on the average growth rate for the next 10 years, and the estimate of long-run growth from a model with a constant intercept for GDP growth. The latter estimate

[27]It is interesting to note that while in our model the innovations to the level of the common factor and its volatility are uncorrelated, the fact that increases in volatility are observed during recessions indicate the presence of negative correlation between the first and second moments, implying negative skewness in the distribution of the common factor. We believe a more explicit model of this feature is an important priority for future research.

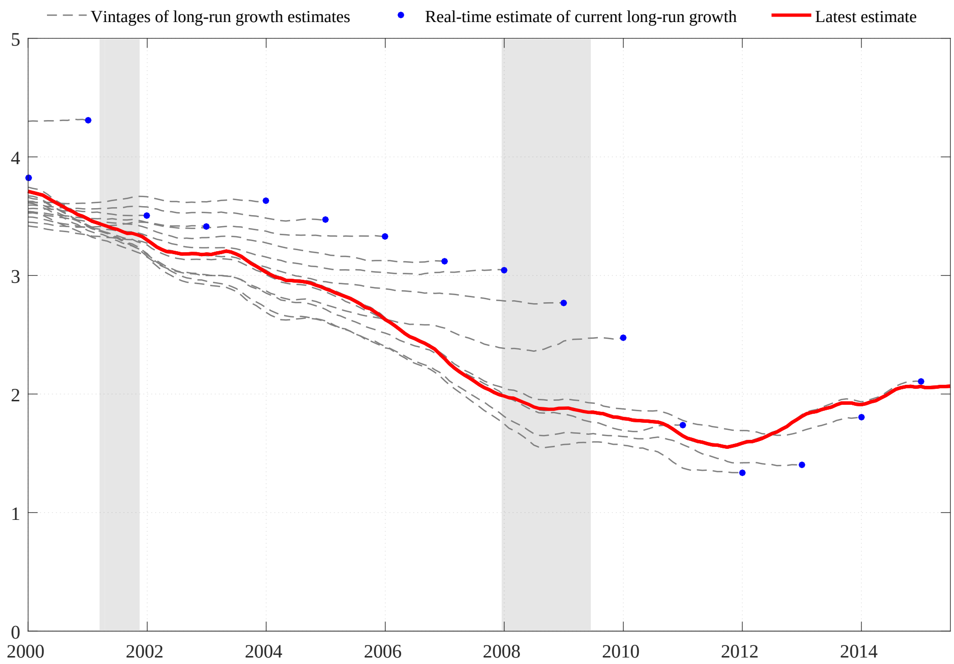
[28]In a few cases new indicators were developed after January 2000. For example, the Markit Manufacturing PMI survey is currently one of the most timely and widely followed indicators, but it started being conducted in 2007. In those cases, we append to the panel, in real time, the vintages of the new indicators as soon sufficient history is available. In the example of the PMI, this is the case since mid-2012. By implication, the number of indicators in our data panel grows when new indicators appear. Full details about the construction of the vintage database are available in Appendix 4.7.6.

Figure 4.3: Long-Run GDP Growth Estimates in Real Time

(a) Evolution of the current assessment of long-run growth



(b) Selected vintages of long-run growth estimates using real-time data



Note: The figure presents results from re-estimating the model using the vintage of data available at each point in time from January 2000 to March 2015. The start of the estimation sample is fixed at Q1:1947. Panel (a) plots the median real-time estimate of current long-run growth over time. This is the locus traced by the end points of all vintages. The blue shaded areas represent the 68th and 90th percentiles. The dashed line is the contemporaneous estimate of the historical average of real GDP growth. The diamonds are the median response to the Philadelphia Fed Livingston Survey of Professional Forecasters on the average growth rate for the next 10 years. Panel (b) displays the median estimate of long-run GDP growth for various vintages of data (dashed gray lines). The estimate of the latest vintage is shown in solid red. Gray shaded areas represent NBER recessions.

is also updated as new information arrives, but weighs all points of the sample equally. Panel (b) displays vintages of the median long-run growth estimate, using information available up to July of each year. The locus traced by the end point of each vintage corresponds to the current real-time estimate of Panel (a).

The evolution of the baseline model's estimate of long-run growth when estimated in real time declines gradually from a peak of about 4% in early 2000 to around 2.5% just after the end of the Great Recession. From this time, the constant estimate shown in panel (a) is always outside of the 90% posterior bands. There is a sharp reassessment of long-run growth around July 2010, coinciding with the publication by the Bureau of Economic Analysis of the annual revisions to the National Accounts, which each year incorporate previously unavailable information for the previous three years. The revisions implied a substantial downgrade, in particular, to the growth of consumption in the first year of the recovery, from 2.5% to 1.6%, and instead allocated much of the growth in GDP during the recovery to inventory accumulation.[29] Reflecting the role of consumption as the most persistent and forward looking component of GDP, the estimate of long-run growth is downgraded sharply. Panel (b) shows how the 2010 revisions in fact trigger a re-interpretation of the years leading to the Great Recession. With the revised information, the bulk of the slowdown in long-run growth is now estimated to have occurred *before* the recession.[30] From 2010 onward, the model predicts a recovery that is extremely slow by historical standards. This is four years before the structural break test detected a statistically significant decline.[31] It is evident from the preceding discussion that revisions to past data by the BEA are an important source of changes to the long-run growth estimate in real time. Since the revision process is not modeled explicitly within the DFM, the in-sample results of Section 4.4.3 do not take into account the uncertainty stemming from future revisions. Interestingly, in the latest part of the sample, the estimate of long-run growth has recovered slightly to about 2% but this has been triggered by improvements in incoming data, rather than revisions to past vintages.

With regards to the SPF, it is noticeable that from 2003 to about 2010, the survey is remarkably similar to the model, but since then, the SPF forecast has continued to drift down only very slowly, standing at 2.5% as of mid-2015. It is noteworthy that, as pointed out by Stanley Fischer in the speech quoted in the introduction, during that period both private and institutional forecasters systematically overestimated growth.

[29] See Appendix 4.7.10 for additional figures on the National Accounts revisions during this period.

[30] Indeed, the (one-sided) filtered estimate based on the latest vintage, which ignores the effect of data revisions, displays a more gradual pattern of decline (see Figure 4.6 in Section 4.7.1 of the Appendix).

[31] A simpler specification that does not use consumption to inform the trend would detect the decline in long-run growth one year later, with additional revisions to past GDP in July 2011.

4.4.5 Implications for Nowcasting GDP

The standard DFM with constant long-run growth and constant volatility has been successfully applied to produce current quarter nowcasts of GDP (see Banbura et al., 2010, for a survey). Using our real-time US database, we carefully evaluate whether the introduction of time-varying long-run growth and SV into the DFM framework also improves the performance of the model along this dimension. We find that over the evaluation window 2000-2015 the model is at least as accurate at point forecasting, and significantly better at density forecasting than the benchmark DFM. We find that most of the improvement in density forecasting comes from correctly assessing the center and the right tail of the distribution, implying that the time-invariant DFM is assigning excessive probability to a strong recovery. In an evaluation subsample spanning the post-recession period, the relative performance of both point and density forecasts improves substantially, coinciding with the significant downward revision of the model's assessment of long-run growth. In fact, ignoring the variation in long-run GDP growth would have resulted in being on average around 1 percentage point too optimistic from 2009 to 2015.[32]

To sum up, the addition of the time-varying components not only provides a tool for decision-makers to update their knowledge about the state of long-run growth in real time. It also brings about a substantial improvement in short-run forecasting performance when the trend is shifting, without worsening the forecasts when the latter is relatively stable. The proposed model therefore provides a robust and timely methodology to track GDP when long-run growth is uncertain.

4.4.6 Inspecting the Role of Data Set Size and Composition

In this paper we argue that the rich multivariate framework of a DFM will facilitate the extraction of the long-run growth component of GDP. The DFM will exploit the cross-sectional dimension, and not just the time series dimension in separating cycle from trend. It is interesting to quantify the advantage that the DFM provides over traditional trend-cycle decompositions, and to investigate the robustness of our main conclusions to alternative datasets of varying size and composition. In order to do so, we consider (1) a bivariate model with GDP and unemployment only (labeled "Okun"), (2) an intermediate model with GDP and the four additional variables often included in the construction of coincident indicators, see Mariano and Murasawa (2003) and Stock and Watson (1989) (labeled "MM03"), (3) our "Baseline" specification with 28 variables, and (4) an "Extended" model that uses disaggregated data for many of the headline series included in the baseline specification, totaling 155 variables.[33] Moreover, in order to investigate the gains associated with imposing

[32] Appendix 4.7.9 provides the full details of the forecast evaluation exercise.

[33] As we argue in Section 4.4.1, the introduction of a large number of disaggregated series, even if related to real activity, is likely to lead to model misspecification whenever the sectoral data are not contemporaneously related. For the extended specification, we consider a solution to this problem which allows to maintain the parsimonious one factor structure. By extending the model to include lags of the factor in the observation equation, each variable can display heterogeneous responses to

additional structure to long-run GDP growth, for the last two specifications we also consider a version of the model that does not impose common long-run growth in GDP and consumption.

The top panel of Table 4.2 reports the mean point-estimates for each specification over selected subsamples.[34] In all cases, the results are consistent with a decline in the long-run growth rate in the last part of the sample. Quantitatively, most specifications are very close to the baseline, with the specifications that impose common long-run growth in GDP and consumption finding an earlier and sharper decline. The exception is the "Okun" specification which instead estimates a smaller increase in the mid 1990's as well as a larger decline in long-run growth in the past decade. It is noteworthy that the mean estimate of the extended specification is very close to that of the baseline.

Table 4.2: RESULT COMPARISON FOR ALTERNATIVE DATA SETS AND SPECIFICATIONS

			Baseline		Extended	
	Okun	*MM03*	*GDP only*	***GDP + C***	*GDP only*	*GDP + C*
Long-run growth						
1947-1972	3.9	3.5	3.6	**3.8**	3.6	3.9
1973-1995	3.2	3.4	3.1	**3.1**	3.2	3.2
1996-2007	2.6	3.2	3.1	**3.1**	3.0	3.1
2008-2015	1.6	2.5	2.4	**1.8**	2.2	1.7
End of Sample	1.3	2.4	2.3	**2.0**	2.1	2.0
Uncertainty: Long run						
Filtered	0.82	0.63	0.64	**0.56**	0.78	0.63
Smoothed	0.44	0.36	0.37	**0.35**	0.44	0.39
Uncertainty: Cycle						
Filtered	2.08	1.47	0.79	**0.76**	0.23	0.23
Smoothed	1.89	1.32	0.62	**0.60**	0.25	0.25

Note: Each column presents the estimation results corresponding to the alternative models (data sets) considered in this section. The upper panel displays the posterior means of the long-run growth rate of real GDP, over selected subsamples. In the lower panel, the posterior uncertainty corresponding to both the long-run growth rate of real GDP, as well as the common factor are displayed. The uncertainty is calculated as an average over the entire sample.

The lower panel of Table 4.2 instead investigates the uncertainty around the mean estimates. The uncertainty around the long-run growth estimate declines as we move from the bivariate to the multivariate specifications, with most of the reduction happening once a handful of variables are included. On the other hand, when the panel is extended to include a large number of disaggregated series, the uncertainty

the common factor, and correlation between idiosyncratic components is reduced. Given that the extended model is heavily parameterized, we follow D'Agostino et al. (2015) in choosing priors that shrink the model towards the contemporaneous-only specification, which is nested in the extended case. Full details and the composition of the data set and the changes to the estimation in case of the extended model are provided in Appendix 4.7.11.

[34]See Figure 4.24 in Appendix 4.7.11 for a comparison of the results of each alternative specification with the baseline results over the entire sample.

increases.[35] While including a few key series, such as the ones in the specification of Mariano and Murasawa (2003) seems to already achieve the bulk of the reduction in uncertainty, it should be taken into account that those variables are available only with a relatively long publication lag, and subject to considerable revisions over time. Our proposed strategy of using an intermediate number of indicators, including the more timely and accurate surveys, is likely to lead to more satisfactory results in a real-time setting. Furthermore, the inclusion of the surveys is helpful in identifying the long-run growth rate, as those variables do not display a time-varying long-run mean by construction.

Overall this exercise highlights that the finding of a substantial decline in the long-run growth rate is confirmed across different specifications that use data sets of varying size and composition. The baseline specification, which uses an intermediate number of series including both hard data and surveys, leads to the lowest uncertainty around the long-run growth estimate, supporting the baseline choice of data set size and composition proposed in Section 4.4.1. Our results have important implications for trend-cycle decompositions of output, which usually include only a few cyclical indicators, generally inflation or variables that are direct inputs to the production function (see e.g. Gordon, 2014a or Reifschneider et al., 2013). As we show, greater precision of the trend component can be achieved by exploiting the common cyclical features of additional macroeconomic variables.[36]

4.5 Decomposing Movements in Long-Run Growth

In this section, we show how our model can be used to decompose the long-run growth rate of output into long-run movements in labor productivity and labor input. By doing this, we exploit the ability of the model to filter away cyclical variation and idiosyncratic noise and obtain clean estimates of underlying long-run trends. We see this exercise as a step towards giving an economically more meaningful interpretation to the movements in long-run real GDP growth detected by our model.

GDP growth is by identity the sum of growth in output per hour and growth in total hours worked. It is therefore possible to split the long-run growth trend in our model into two orthogonal components such that this identity is satisfied in the long run. Here we make use of our flexible definition of $\mathbf{c}_t$ in equation (4.2). In particular, ordering the growth rates of real GDP, real consumption and total hours as the first three variables in $\mathbf{y}_t$, we define

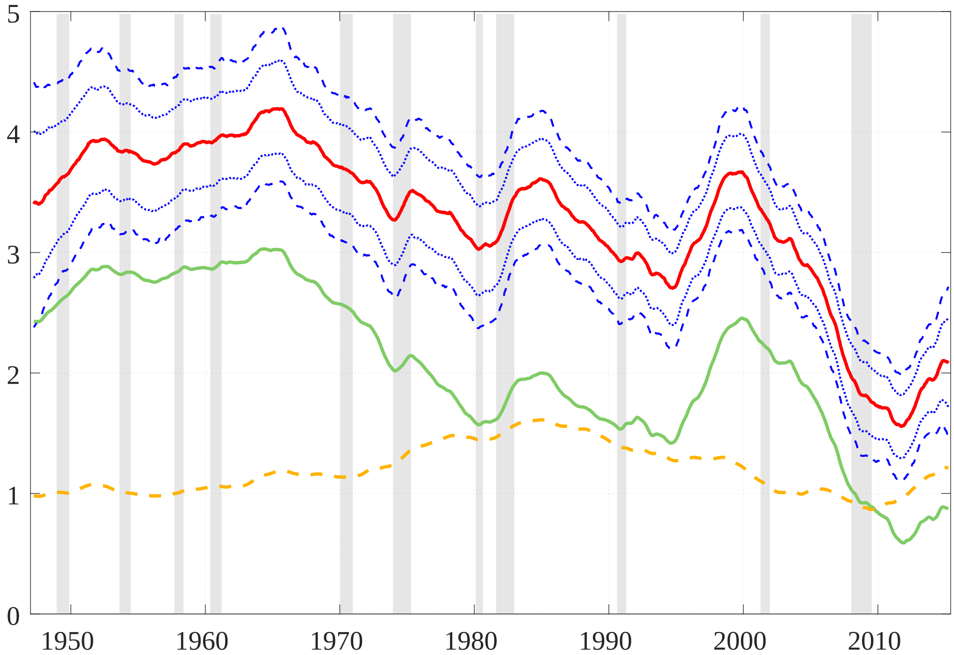
$$\mathbf{a}_t = \begin{bmatrix} z_t \\ h_t \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, \tag{4.10}$$

[35]We conjecture that as many more variables are added, the fit of the common factor to the cyclical component of GDP worsens. As a consequence, some cyclical variation of GDP spills over to the estimate of the long-run component. The uncertainty around the common factor, on the other hand, continues to decline.

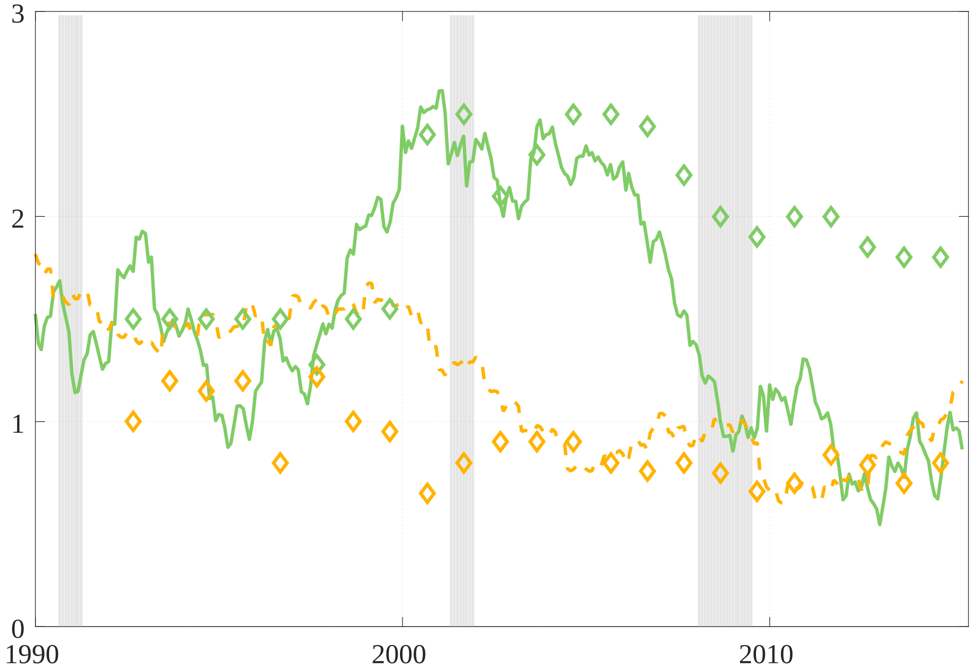
[36]Basistha and Startz (2008) make a similar point, arguing that the inclusion of indicators that are informative about common cycles can help reduce the uncertainty around Kalman filter estimates of the long-run rate of unemployment (NAIRU).

Figure 4.4: Decomposition of Long-run US Output Growth

(a) Posterior median estimates of decomposition



(b) Filtered estimates of long-run growth components



Note: Panel (a) plots the posterior median (solid red), together with the 68% and 90% (doted and dashed blue) posterior credible intervals of long-run GDP growth and the posterior median of both long-run labor productivity growth and long-run total hours growth (solid green and dashed orange). Panel (b) plots the filtered estimates of these two components, i.e. $\hat{z}_{t|t}$ and $\hat{h}_{t|t}$, since 1990. For comparison, the corresponding forecasts from the SPF are plotted. The SPF forecast for total hours is obtained as the difference between the forecasts for real GDP and labor productivity.

so that the model is specified with two time-varying components, the first of which loads output and consumption but not hours, and the second loads all three series. The first component is then by construction the long-run growth rate of labor productivity, while the second one captures low-frequency movements in labor input independent of productivity.[37] Given the relation in (4.10), the two components add up to the time-varying intercept in the baseline specification, i.e. $g_t = z_t + h_t$.[38] It follows from standard growth theory that our estimate of the long-run growth rate of labor productivity will capture both technological factors and other factors, such as capital deepening and labor quality.[39]

Figure 4.4 presents the results of the decomposition exercise for the US. Panel (a) plots the median posterior estimate of long-run real GDP growth and its labor productivity and total hours components. The posterior bands for long-run real GDP growth are included. The time series evolution conforms very closely to the narrative of Fernald (2014), with a pronounced boom in labor productivity in the mid-1990's and a subsequent fall in the 2000's clearly visible. The decline in the 2000's is relatively sudden while the 1970's slowdown appears as a more gradual phenomenon starting in the late 1960's. Furthermore, the results reveal that during the 1970's and 1980's the impact of the productivity slowdown on output growth was partly masked by a secular increase in hours, probably reflecting increases in the working-age population as well as labor force participation (see e.g. Goldin, 2006). Focusing on the period since 2000, labor productivity accounts for almost the entire decline.[40] This contrasts explanations by which slow labor force growth has been a drag on GDP growth. When taking away the cyclical component of hours and focusing solely on its long-run component, the contribution of hours has, if anything, accelerated since the Great Recession. Panel (b) presents the filtered estimates of the two components, i.e. the output of the Kalman Filter which uses data only up to each point in time. For comparison, the corresponding SPF forecasts are included. Most notably, this plot reveals that starting around 2005 a relatively sharp revision to labor productivity drives the decline in long-run output growth.[41] Interestingly,

[37] z_t and h_t jointly follow random walks with diagonal covariance matrix as defined by equation (4.7). Restricting the covariance matrix is not necessary for estimation, but imposing it allows us to interpret the innovations to the trends as exogenous shocks to the long-run growth rates of the variables. The hours trend is therefore interpreted as those low-frequency movements in hours which are uncorrelated with labor productivity. Allowing for a full covariance matrix would yield trends that are linear combinations of the current ones, but would lack a clear economic interpretation.

[38] Since z_t and h_t are independent and add up to g_t, we set the prior on the scale of their variances to half of the one set in Section 4.4.2 on g_t. In addition, note that the cyclical movement in labor productivity is given by $(1 - \lambda_3)f_t$.

[39] Further decomposing z_t into technology and non-technology movements requires additional information to separately identify these components. One possibility, which we explore in Appendix 4.7.12, is to use an independent measure of TFP to isolate technological factors. Note, however, that reliable data on capital input, labor quality, or estimates of TFP are not available in real time, making the focus on long-run labor productivity more appealing in a real-time setting.

[40] In Appendix 4.7.12 we extend the analysis to decompose the labor productivity trend into long-run TFP and non-technological forces. We find that TFP accounts for virtually all of the slowdown.

[41] In an additional figure, provided in Section 4.7.1 of the Appendix, we plot 5,000 draws from the joint posterior distribution of the variances of the innovations to the labor productivity and hours components. This analysis confirms the conclusion from the discussion here that changes in

the professional forecasters have been very slow in incorporating the productivity slowdown into their long-run forecasts. This delay explains their persistent overestimation of GDP growth since the recession.

It is interesting to compare the results of our decomposition exercise to similar approaches in the literature, in particular Gordon (2010, 2014a) and Reifschneider et al. (2013). Like us, they specify a state space model with a common cyclical component and use the 'output identity' to decompose the long-run growth rate of GDP into underlying drivers. A key difference resides in the Bayesian estimation of the model, which enables us to impose a conservative prior on the variance of the long-run growth component that helps avoiding over-fitting the data. Furthermore, the inclusion of SV in the cyclical component helps to prevent unusually large cyclical movements from contaminating the long-run estimate. Another important difference is that we use a larger amount of information, including key cyclical indicators like industrial production, sales, and business surveys, which are generally not included in a production function approach. This allows us to retrieve a timely and precise estimate of the cyclical component and, as a consequence, to reduce the uncertainty that is inherent to any trend-cycle decomposition of the data, as discussed in Section 4.4.6. As a result, we obtain a substantially less pessimistic estimate of the long-run growth of GDP than these studies in the latest part of the sample. For instance, Gordon (2014a) reports a long-run GDP growth estimate below 1% for the end of the sample, whereas our median estimate stands at around 2%.[42]

4.5.1 International Evidence

To gain an international perspective on our results, we estimate the DFM for the other G7 economies and perform the decomposition exercise for each of them.[43] The median posterior estimates of the labor productivity and labor input trends are displayed in Figure 4.5. Labor productivity, displayed in Panel (a), plays again the key role in determining movements in long-run growth. In the Western European economies and Japan, the elevated growth rates of labor productivity prior to the 1970's reflect the rebuilding of the capital stock from the destruction from World War II, and ended as these economies converged towards US levels of output per capita. The labor productivity profile of Canada broadly follows that of the US, with a slowdown in the 1970's and a temporary mild boom during the late 1990's. Interestingly, this acceleration in the 1990's did not occur in Western Europe and

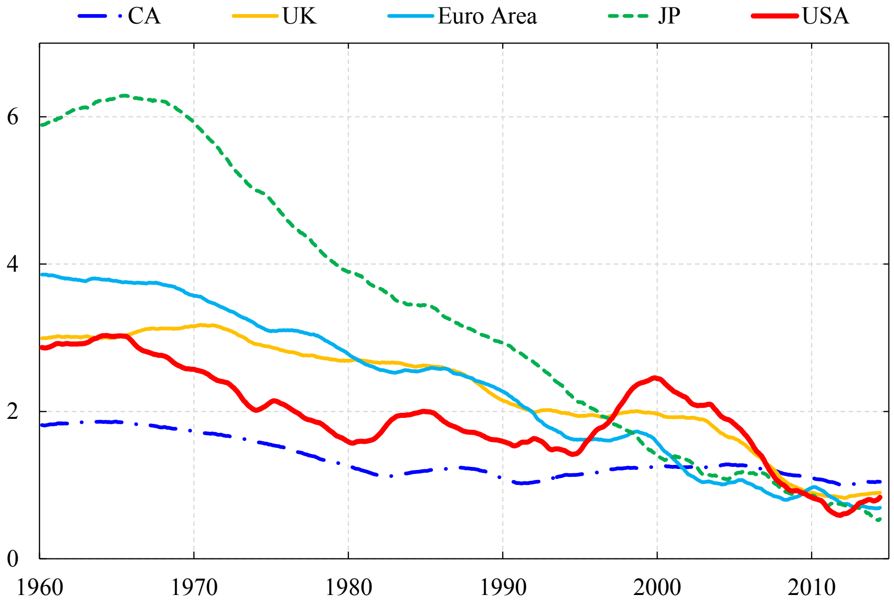
labor productivity, rather than in labor input, are the key driver of low frequency movements in real GDP growth.

[42]The results for a bivariate model of GDP and unemployment, which we have discussed in Section 4.4.6 show that the current long-run growth estimate is 1.3%, close to Gordon (2014a).

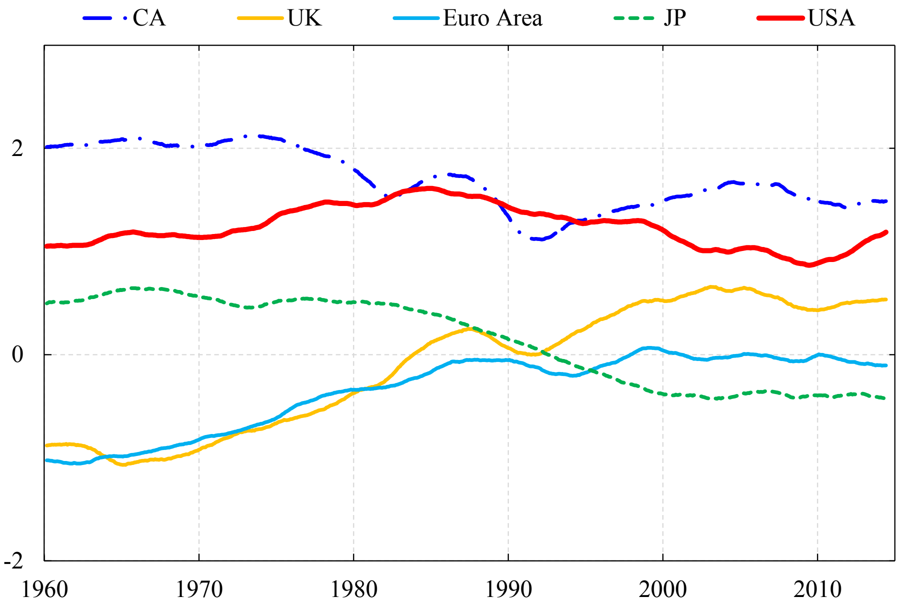
[43]Details on the specific data series used for each country are available in Appendix 4.7.6. For hours, we again follow the methodology of Ohanian and Raffo (2012). In the particular case of the UK, the quarterly series for hours displays a drastic change in its stochastic properties in the early 1990's owing to a methodological change in the construction by the ONS, as confirmed by the ONS LFS manual. We address this issue by using directly the annual series from the TED, which requires an appropriate extension of equation (4.9) to annual variables (see Banbura et al. 2012). To avoid weak identification of h_t for the UK, we truncate our prior on its variance to discard values which are larger than twice the maximum posterior draw of the case of the other countries.

Figure 4.5: Decomposition for Other Advanced Economies

(a) Long-run Labor Productivity



(b) Long-run Labor Input



Note: Panel (a) displays the posterior median of long-run labor productivity across advanced economies. Panel (b) plots the corresponding estimates of long-run total hours worked. In both panels, 'Euro Area' represents a weighted average of Germany, Italy and France.

Japan.[44] The UK displays a decline in labor productivity similar to the US. This "productivity puzzle" has been debated extensively in the UK (see e.g. Pessoa and Van Reenen, 2014). It is interesting to note that the two countries which experienced a more severe financial crisis, the US and the UK, appear to be the ones with greatest declines in productivity since the early 2000's, similar to the evidence documented in Reinhart and Rogoff (2009).

Panel (b) displays the movements in long-run hours worked identified by equation (4.10). The contribution of this component to overall long-run output growth varies considerably across countries. However, within each country it is more stable over time than the productivity component, which is in line with our findings for the US. Indeed, the extracted long-run trend in total hours includes various potentially offsetting forces that can lead to changes in long-run output growth. In any case, the results of our decomposition exercise indicate that after using the DFM to remove business-cycle variation in hours and output, the decline in long-run GDP growth that has been observed in the advanced economies since the early 2000's is entirely accounted for by a decline in the labor productivity trend. Finally, it is interesting to note that for the countries in the sample long-run productivity growth appears to converge in the cross section, while there is no evidence of convergence in the long-run growth of hours.[45]

4.6 Conclusion

The sluggish recovery from the Great Recession has raised the question whether the long-run growth rate of US real GDP is now lower than it has been on average over the postwar period. We have presented a DFM that allows for both changes in long-run GDP growth and stochastic volatility. Estimating the model with Bayesian methods, we provide evidence that long-run growth of US GDP displays a gradual decline after the turn of the century, moving from its peak of 3.5% to about 2% in 2015. Using real-time vintages of data we demonstrate the model's ability to track GDP in a timely manner. By the summer of 2010 the model would have concluded that a significant decline in long-run growth was behind the slow recovery, therefore substantially improving the real-time tracking of GDP by taking into account the uncertainty surrounding long-run growth. Finally, we discuss the drivers of movements in long-run output growth through the lens of our model by decomposing it into the long-run growth rates of labor productivity and labor input. Using data for both the US and other advanced economies our model points to a global slowdown in labor productivity as the main driver of weak growth in recent years, extending the narrative of Fernald (2014) to other economies. Studying the deep causes of the secular decline in growth is an important priority for future research.

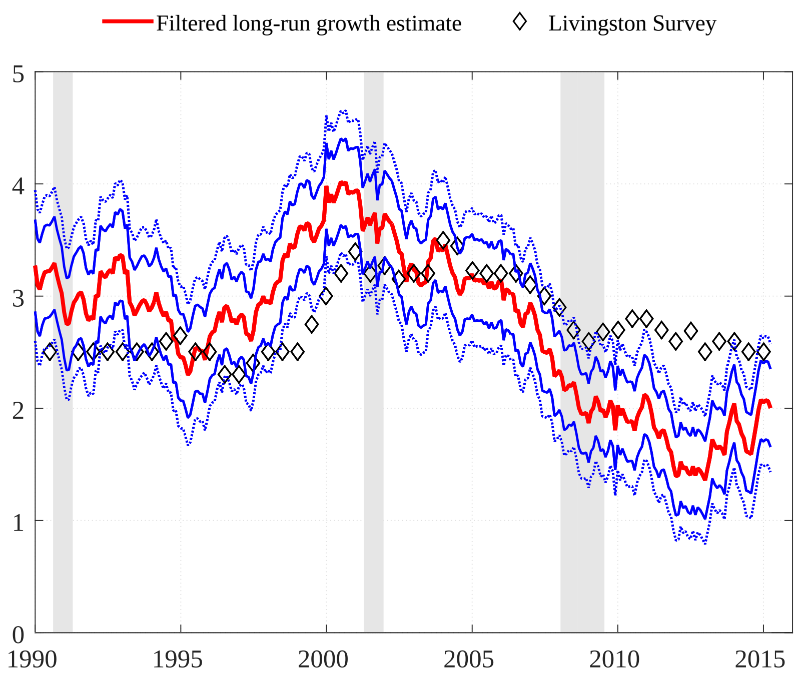
[44] On the lost decade in Japan, see Hayashi and Prescott (2002). Gordon (2004) examines the absence of the IT boom in Europe.

[45] Similar evidence for emerging economies has been recently presented by Pritchett and Summers (2014). Their evidence refers to convergence of overall GDP growth rates, whereas ours indicates that convergence in productivity growth appears to be the dominant source of convergence.

4.7 Appendices

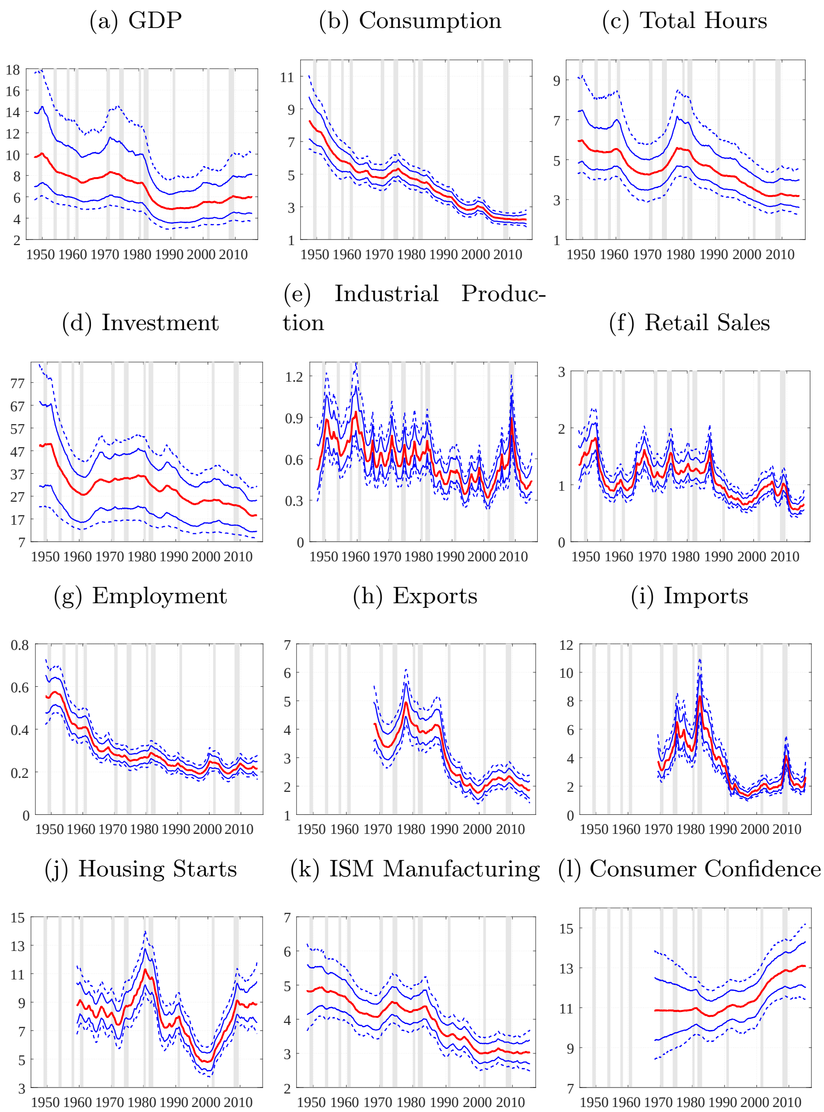
4.7.1 Additional Figures

Figure 4.6: Filtered estimate of long-run growth



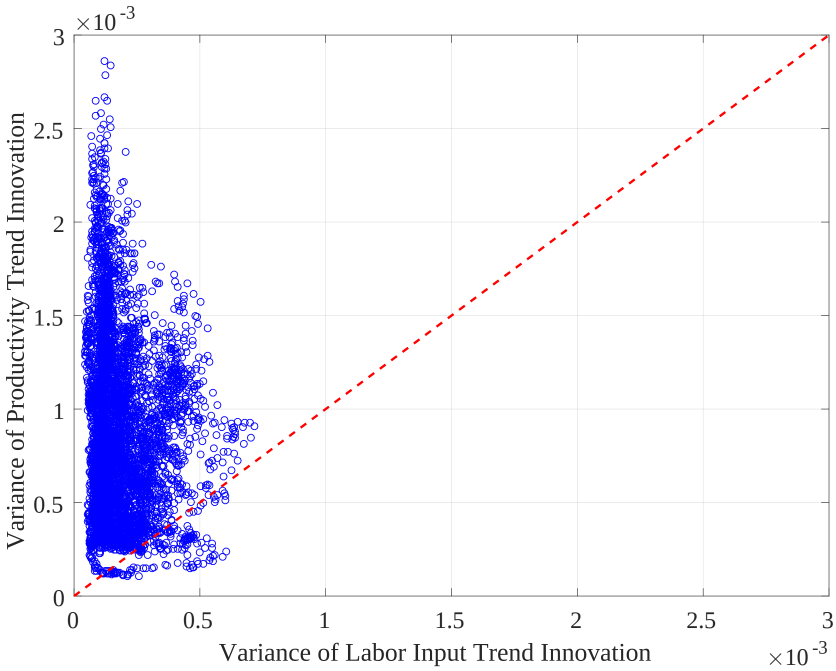
Note: The solid red line is the filtered estimate of the long-run GDP growth rate, $\hat{g}_{t|t}$, using the vintage of National Accounts available as of March 2015. The solid and dotted blue lines capture the corresponding 68% and 90% posterior bands. The black diamonds represent the real-time mean forecast from the Livingston Survey of Professional Forecasters of the average GDP growth rate for the subsequent 10 years.

Figure 4.7: Stochastic Volatility of Selected Idiosyncratic Components



Note: Each panel presents the median (solid red), the 68% and the 90% (solid and dashed blue) posterior credible intervals of the volatility of the idiosyncratic component of selected variables. Shaded areas represent NBER recessions. Similar charts for other variables are available upon request.

Figure 4.8: Joint Posterior Distribution of Growth Component Innovation Variances



Note: The figure plots 5,000 draws of the joint posterior distribution of the variances of innovations to the labor productivity and hours component. The dashed red line is the 45-degree line. Under the equal-variance prior the draws would be equally distributed above and below this line. The fact that the bulk of draws lie above indicates that changes in long-run labor productivity drive the variation in long-run output.

4.7.2 Full Results of Structural Break Tests

Nyblom Test

Table 4.3 reports the result for the Nyblom (1989) test applied to US real GDP growth, as described in Hansen (1992). The sample starts is 1947:Q2. The specification is $y_t = \mu + \rho_1 y_{t-1} + \rho_2 y_{t-1} + \sigma\epsilon_t$, where y_t is real GDP growth. For each parameter of the specification, the null hypothesis is that the respective parameter is constant.

Table 4.3:
TEST RESULTS OF NYBLOM TEST

	L_c	
	AR(1)	AR(2)
μ	0.518*	0.473*
ρ_1	0.367	0.331
ρ_2		0.094
σ^2	0.843***	0.838***
Joint L_c	2.145***	2.294***

Note: Results are obtained using Nyblom's L test as described in Hansen (1992). *, ** and *** indicate significance at the 10%, 5% and 1% level.

Bai and Perron Test

Table 4.4 reports the result for the Bai and Perron (1998) test applied to US real GDP growth for the sample starting in 1947:Q2. We apply the $SupF_T(k)$ test for the null hypothesis of no break against the alternatives of $k = 1, 2$, or 3 breaks. Secondly, the test $SupF_T(k+1|k)$ tests the null of k breaks against the alternative of $k+1$ breaks. Finally, the $U_d max$ statistic tests the null of absence of break against the alternative of an unknown number of breaks. The null hypothesis of no breaks is rejected against the alternative of one break at the 10% level. The null is not rejected against the alternative of two or three breaks. Furthermore, the null hypothesis of one break against two breaks, or the null of only two against three breaks is not rejected. The final test confirms the conclusion that there is some evidence in favor of at least one break, with the null rejected against an unknown number of breaks at the 10% level. The most likely break is identified to have happened in the second quarter of 2000.

Table 4.4:
TEST RESULTS OF BAI-PERRON TEST

	Sample 1947-2015
	$SupF_T(k)$
$k = 1$	8.379*
	[2000:Q2]
$k = 2$	4.194
	[1968:Q2; 2000:Q2]
$k = 3$	4.337
	[1969:Q1; 1982:Q4; 2000:Q2]
	$SupF_T(k\|k-1)$
$k = 2$	1.109
$k = 3$	2.398
$U_d max$	8.379*

Note: Results are obtained using the Bai and Perron (1998) methodology. Dates in square brackets are the most likely break date(s) for each of the specifications. * indicates significance at the 10% level.

4.7.3 Monte Carlo Evidence

Setup for Monte Carlo simulations

To assess the performance of our model in the presence of potentially relevant types of misspecification, we carry out a variety of Monte Carlo experiments. In each experiment, we simulate a large number of data sets which are generated from the model under known parameter values, and estimate our model repeatedly over these data sets. This appendix presents the results for two sets of such experiments, which are designed to explore the robustness of crucial assumptions made in the paper.

- In Section 4.7.3 we examine whether the random walk assumption for the time-varying parameters is robust to a different type of structural change. In particular, we verify how the model performs if the underlying long-run growth rate of GDP features one or multiple discrete breaks rather than gradual change. We also estimate our baseline model on data which is generated with a constant instead of a time-varying long-run growth rate of real GDP growth. Furthermore, we repeat this type of experiment for discrete breaks rather than gradual change in the volatilities of both the common factor and the idiosyncratic terms.

- In Section 4.7.3 we explore the robustness of our model to the presence of (unmodeled) change in the long-run growth rate of other series. We entertain the possibility that such unmodeled trends are either independent of the change in the long-run growth real GDP growth or that some series share the trend of GDP. We also verify robustness to both of these types of misspecification simultaneously.

We aim to ensure a realistic environment for the correctly specified parts of the model. In particular, we set the values of the parameters to their estimated posterior median of the US results. We then take draws for the random disturbances and generate a sample of the vector of 28 observables using equations (4.1) to (4.7), and generate 800 periods of data, which corresponds to the monthly sample size in our US application.[46] The four quarterly series are generated by simulating the underlying monthly series and then introducing missing observations by (backwards) applying the polynomial in equation (4.9). We then estimate the model using the settings described in the paper. The number of simulations (repeatedly drawn samples) per given experiment is set to 100.[47]

[46]While we argue in the paper that the random walk assumption for the estimation of the time-varying parameters is innocuous, it can be problematic to simulate data from parameters that follow random walks. Although we would like the parameters to drift in a non-stationary fashion, i.e. to generate realistic patterns of time-varying volatility, data sets generated from "explosive" processes feature unrealistic properties. To address this issue in the Monte Carlo simulations we discard and re-generate random walks when they drift across a fix threshold. For example, we do not allow the range of (demeaned) time-varying intercept of a given series to exceed the range of its cyclical component.

[47]In certain cases, convergence of the algorithm takes longer in the presence of misspecification, which required us to increase the number of draws of the Gibbs sampler, and thus limited the amount of repetitions that was feasible for a given experiment.

Results: Sensitivity of random walk specification

The goal of this first set of Monte Carlo experiments is to explore the sensitivity of our modeling choice with respect to the random walk specification of the time-varying parameters. The details about how we justify this modeling assumption can be found in Section 4.3.1 of the paper. In particular, we aim here to verify whether the model is robust in a context in which there are changes in the long-run growth rate of real GDP growth and in the volatility of business cycles, but these changes occur as discrete breaks rather then as gradual change. Figures 4.9 to 4.12 present the results of four Monte Carlo experiments.

In the first experiment, the simulated counterpart of real GDP growth features a mean growth rate that is constant but subject to a level shift in the middle of the sample. In Figure 4.9, panels (a) and (b), we plot the actual growth rate underlying the data-generating process together with one and two standard deviation percentiles of the 100 simulations of the posterior median, both for the filtered and smoothed estimate. It is reassuring to see that the random walk process "learns" relatively quickly about the underlying change, even in the case of a discrete jump. Panel (c) displays the true, together with the posterior estimate of the common factor for one of the 100 Monte Carlo draws. Panel (d) provides a scatter plot of the true vs. estimated stochastic volatilities. Both pictures show that the models performs well at capturing the simulated objects.

In the second experiment, we repeat the same exercise in the presence of two discrete breaks in the real GDP growth rate. The results are visible in Figure 4.10, which tells a very similar story to the first experiment. We omit panels for factor and the stochastic volatility estimates, as they are very similar to the first experiment.

In the third experiment, we verify the consequences of estimating our model in an in environment in which the parameters which we specify as time-varying are in fact constant in the data-generating process. The results, displayed in Figure 4.11, confirm that the random walk assumption appears to be entirely innocuous in this setting. Both the long-run GDP growth rate (smoothed and filtered), as well as the volatility of the factor are estimated to be constant, with relatively high precision. In addition, similar to the first experiment, the estimate of the common factor is very precise.

Finally, in the fourth experiment, we again keep the long-run growth rate of real GDP constant but this time introduce a discrete shift in the volatilities of both the common factor and the idiosyncratic terms of all series in the middle of generated data sample. Reassuringly, the shift in the volatilities is well captures in the estimation and does not spill over to the estimate of the long-run growth rate of real GDP.

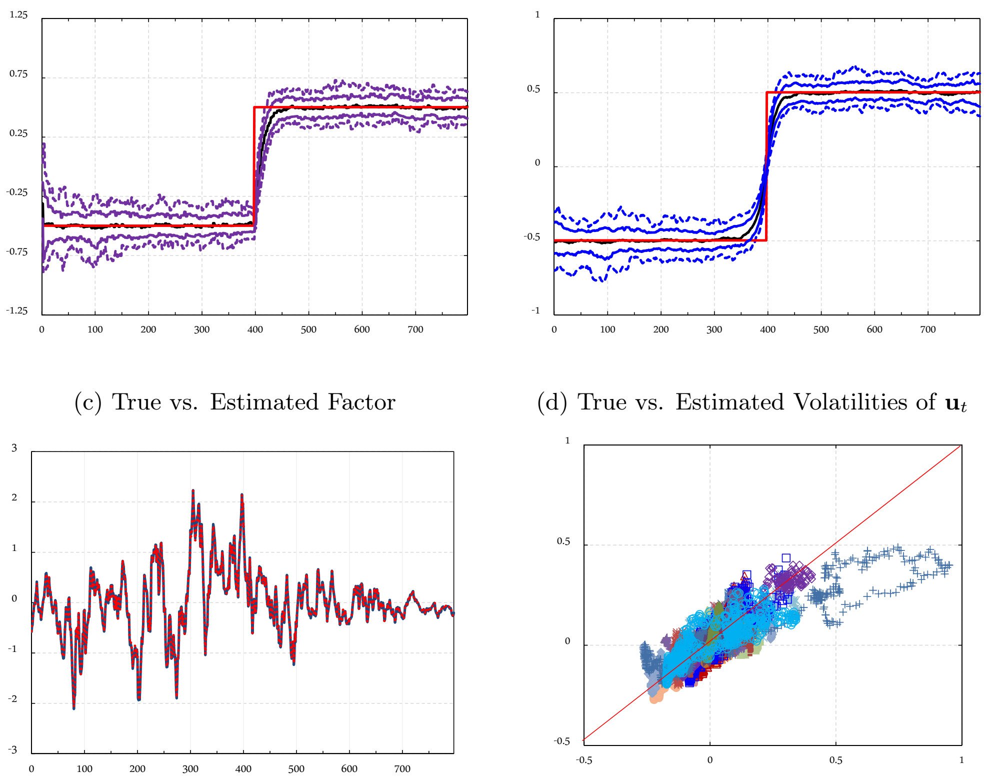
In conclusion from these experiments, the random walk assumption appears to be flexible enough to accommodate structural change that occurs in discrete steps rather than gradually. This underpins our conclusions about the apparent gradual changes in the long-run growth of the US economy described in the paper.

Figure 4.9: Simulation Results I

Data-generating process (DGP) with one discrete break in long-run real GDP growth

(a) True vs. Estimated Trend (Filtered) | (b) True vs. Estimated Trend (Smoothed)

(c) True vs. Estimated Factor | (d) True vs. Estimated Volatilities of $\mathbf{u}_t$

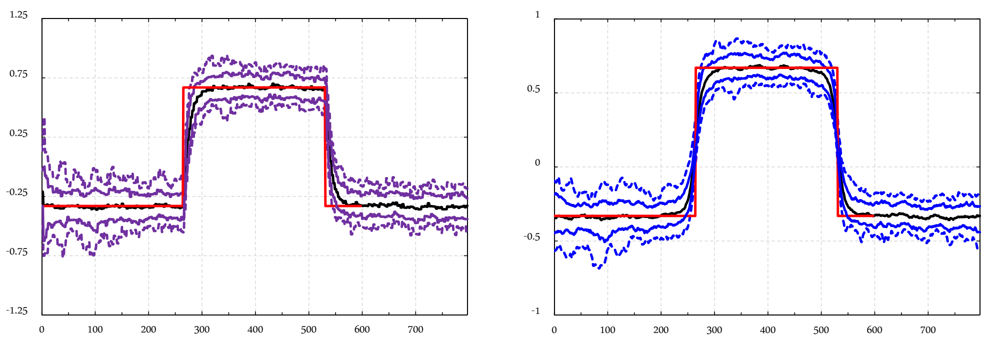


Note: The DGP features a discrete break in the trend of GDP growth occurring in the middle of the sample, as well as stochastic volatility. The sample size is $n = 28$ and $T = 800$, which mimics our US data set. The estimation procedure is the fully specified model as defined by equations (4.1)-(4.7) in the text. We carry out 100 simulations drawing from the DGP. Panel (a) presents the long-run growth component as estimated by the Kalman filter, plotted against the actual long-run growth rate generated from the DGP. The corresponding figure for the smoothed estimate is given in panel (b). In both panels, the median (black) as well the 68th (solid) and 90th (dashed) percentile of the 100 simulated outcomes are shown in blue/purple. Panel (c) displays the factor generated by the the DGP (red) and its smoothed estimate (blue) for one draw. Panel (d) provides evidence on the accuracy of the estimation of the SV of the idiosyncratic terms, by plotting the volatilities from the DGP against the estimates for the 24 monthly indicators. Both are normalized by subtracting the average volatility.

Figure 4.10: Simulation Results II

Data-generating process (DGP) with two discrete breaks in in long-run real GDP growth

(a) True vs. Estimated Trend (Filtered) (b) True vs. Estimated Trend (Smoothed)

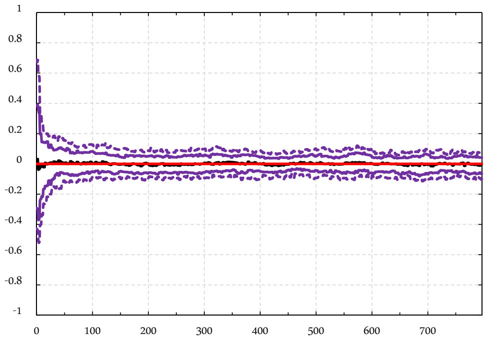


Note: The simulation setup is equivalent to the one in Figure 4.9 but features *two* discrete breaks in the trend at 1/3 and 2/3 of the sample. Again, we show the filtered as well as the smoothed trend median estimates and the corresponding 68th and 90th percentiles of the 100 simulated estimates of these objects. Panels (c) and (d) are omitted as they are very similar to Figure 4.9.

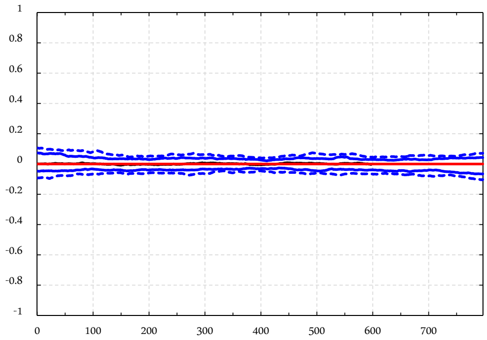
Figure 4.11: Simulation Results III

Data-generating process (DGP) without in changes long-run real GDP growth and without SV

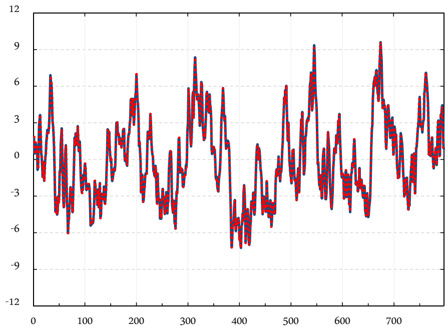
(a) True vs. Estimated Trend (Filtered)



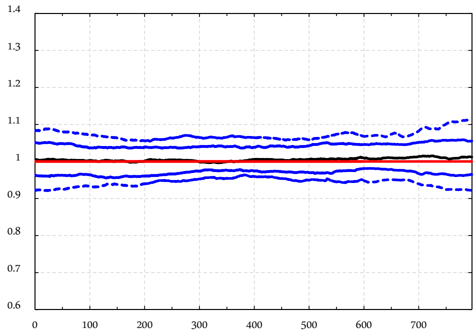
(b) True vs. Estimated Trend (Smoothed)



(c) True vs. Estimated Factor



(d) True vs. Estimated Volatility of Factor

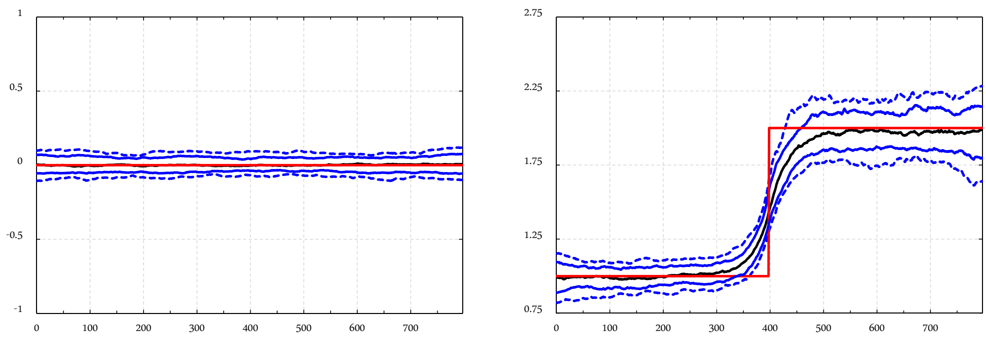


Note: The DGP is the baseline model without trend in GDP growth and without stochastic volatility. The estimation procedure is the fully specified model as explained in the description of Figure 4.9. Again, we plot the filtered and smoothed median estimates of the long-run growth rate with 68th and 90th percentiles of the 100 simulated estimates in panels (a) and (b). Panel (c) presents a comparison of the estimated factor and its DGP counterpart for one Monte Carlo draw. Panel (d) in similar to (b), but for the volatility of the common factor.

Figure 4.12: Simulation Results IV

Data-generating process (DGP) with discrete break in volatility

(a) True vs. Estimated Trend (Smoothed)　　(b) True vs. Estimated Volatility of Factor



Note: The DGP does not feature any changes in the trend of GDP growth, but one discrete break in the volatility of the common factor. As in Figures 4.9-4.11, the estimation procedure is based on the fully specified mode. Panel (a) displays the smoothed posterior median estimate of the trend component of GDP growth, with 68th and 90th percentiles of the 100 simulations shown as solid and dashed blue lines, respectively. Panel (b) displays the posterior median estimate of the volatility of the common factor (black), with the corresponding percentiles.

Results: Sensitivity to confounding time-variation

Our model can flexibly accommodate time-varying intercepts in all or a subset of the series contained in our data panel. Given our interest in tracking real GDP growth, we restrict our baseline model to feature a trend in GDP only (shared by consumption) and argue that such unmodeled time-variation is picked up by the idiosyncratic components, which we allow to be persistent. Details about this discussion are contained in Section 4.3.2 of the paper. The goal of this second set of Monte Carlo experiments is to verify how robust our model is in a setting where time-varying intercepts are indeed present in the data-generating process but not modeled explicitly in the estimation. Figures 4.13 to 4.15 present the results of three Monte Carlo experiments in which such "confounding trends" are added when generating the data.[48]

In the first experiment, the misspecification arises from the fact that our model explicitly specifies a time-varying mean in the GDP equation only, while the data is generated such that the first 18 series of the panel all feature independent non-stationary means.[49] Figure 4.13 presents the estimation results in this setup. Panel (a) shows the percentiles of the deviations of the estimated from the actual real GDP growth rates over the 100 simulations (repeated draws from the DGP). The percentiles are centered relatively tightly around zero, meaning that the trend estimates with 68 and 90% smallest deviations are relatively similar to the original trend process. To illustrate this further, panels (b), (c) and (d) display more detailed results for one of the 100 Monte Carlo simulations, labeled "Median Simulation". This is selected by ordering the outcomes of all repeated samples by the distance of squared deviation of the estimated from the simulated GDP trend and then selecting the median. This essentially means that 49% of the simulations had larger, and 50% smaller deviations than the simulation displayed. The panels plot actual against estimated (black/red) long-run real GDP growth rate, common factor and factor volatility, respectively. In the case of the long-run growth rate the posterior credible intervals are added in blue. These results reveal that in a typical (median) outcome for this type of specification, the model performs well at capturing these objects. Most importantly, the "true" long-run growth rate is contained within the posterior bands throughout the entire estimation sample.

In the second experiment, the data-generating process features a single time-varying mean which is present in the first 6 series, whereas we still only specify it in the first series for the estimation.[50] The results for this experiment are shown in Figure 4.14. The panels here are similar to Figure 4.13. While the deviations

[48]For simplicity we assume that the estimated model in this section is the one with a trend in GDP only, i.e. $\mathbf{B} = 1$.

[49]Formally, in the DGP $dim(\mathbf{a}_t) = 18$ and $\mathbf{B} = \mathbf{I}_{18}$, while the model for estimation is specified by $\mathbf{a}_t = g_t$ and $\mathbf{B} = 1$. We assume that the remaining 10 of the 28 series are stationary, which mimics the presence of the surveys in our data set.

[50]In our notation this means that in the DGP we have $\mathbf{a}_t = g_t$ and $\mathbf{B} = \mathbf{1}_{6\times1}$, while the model for estimation is specified by $\mathbf{a}_t = g_t$ and $\mathbf{B} = 1$. We choose 6 series so that both quarterly and monthly variables are affected by the misspecification.

in panel (a) are slightly larger than for the previous figure, indicating that common unmodeled trends are somewhat more challenging to pick up than independent ones, the overall message remains the qualitatively similar. In particular, the results for the "Median Experiment", displayed in panels (b) to (d), are reassuring in that the estimate tracks their data counterpart closely.

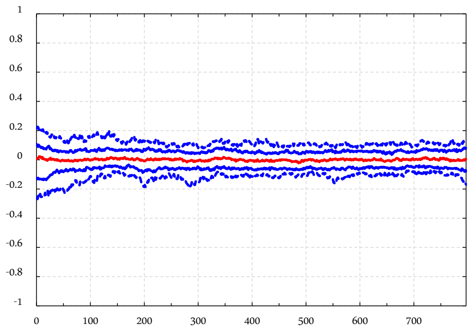
The third experiment introduces both types of misspecification simultaneously, i.e. independent time-varying means in series 1-18 and an additional shared time-varying component in series 1-6. The results are presented in Figure 4.15. The take-aways are similar to the previous figures, even in the presence of this heavy type of misspecification.

Overall, these simulation experiments confirm our intuition that the estimate of the time-varying mean of interest is not affected by low frequency movements present in other series that are not explicitly modeled. Despite the extremely unfavorable assumption of a large amount of additional time-variation, the long-run growth rate of real GDP is tracked very well in all settings considered.

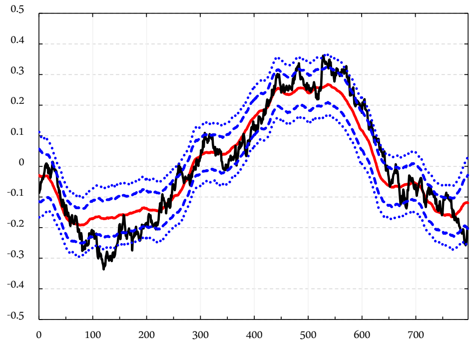
Figure 4.13: Simulation Results V

Data-generating process (DGP) with independent unmodeled trends in other series

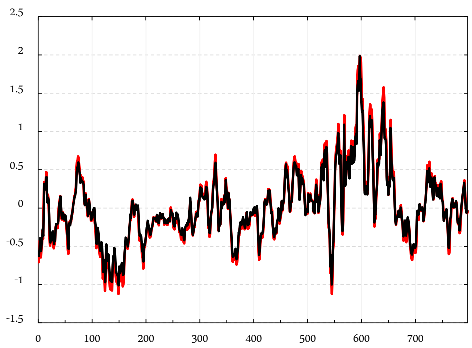
(a) True vs. Est. Trend - Deviation Percentiles



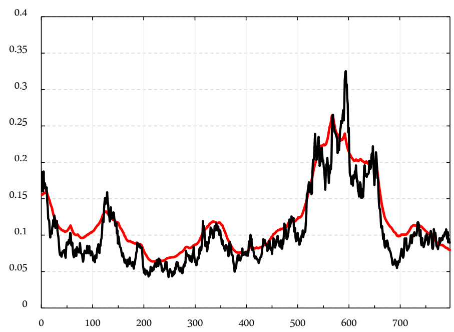
(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)



(d) True vs. Est. Vol (Median Simulation)

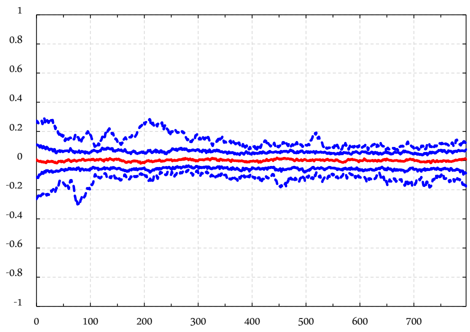


Note: The DGP features independent time-varying means in series 1-18. The sample size is $n = 28$ and $T = 800$, which mimics our US data set. The estimation procedure is the fully specified model as defined by equations (4.1)-(4.7) in the text, with a time-varying mean only specified for the real GDP growth equation. We carry out a Monte Carlo simulation with 100 samples repeatedly drawn from the DGP. Panel (a) presents the median (red), as well as the 68 and 90% bands (blue) of the deviation of the estimated long-run growth rate from its actual data counterpart over 100 simulated outcomes. Panel (b) shows the true (black) together with the posterior median estimate (red) of the long-run growth rate of real GDP. The 68% (solid blue) and 90% (dashed blue) posterior credible intervals are also plotted. Panels (c) and (d) plot the median estimate (red) against true (black) common factor and its stochastic volatility.

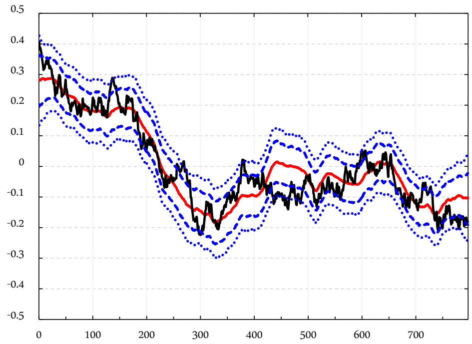
Figure 4.14: Simulation Results VI

Data-generating process (DGP) with shared unmodeled trends in other series

(a) True vs. Est. Trend - Deviation Percentiles

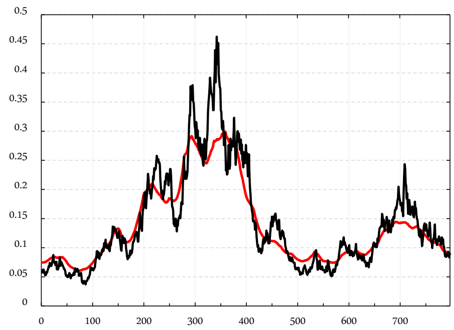


(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)

(d) True vs. Est. Vol (Median Simulation)

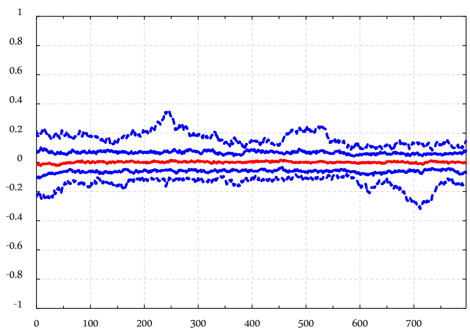


Note: The DGP features a common time-varying in series 1-6, while the estimation specifies this stochastic trend only in the equation for real GDP growth. The rest of the setup of the simulations, as well as the structure of the panels are similar to Figure 4.13.

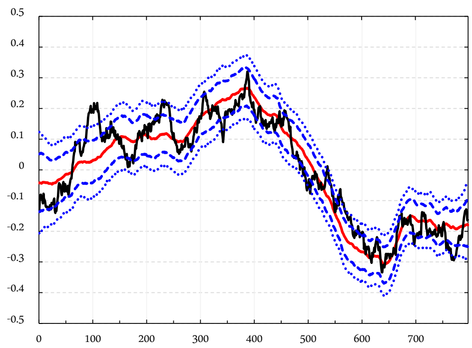
Figure 4.15: Simulation Results VII

Data-generating process (DGP) with both independent and shared unmodeled trends in other series

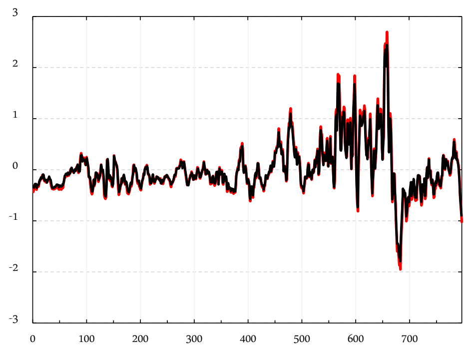
(a) True vs. Est. Trend - Deviation Percentiles



(b) True vs. Est. Trend (Median Simulation)



(c) True vs. Est. Factor (Median Simulation)



(d) True vs. Est. Vol (Median Simulation)

The DGP features both independent time-varying components in series 1-18 as well as a common time-varying in series 1-6, while the estimation specifies a stochastic trend only in the equation for real GDP growth. The rest of the setup of the simulations, as well as the structure of the panels are similar to Figure 4.13

4.7.4 Details on Estimation Procedure

Construction of the State Space System

For expositional clarity, we focus on the baseline case with $\mathbf{B} = 1$ and $\mathbf{a}_t = a_t$ here, so that $m = r = 1$. Recall that in our main specification we choose the order of the polynomials in equations (4.3) and (4.4) to be $p = 2$ and $q = 2$, respectively. Let the $n \times 1$ vector $\tilde{\mathbf{y}}_t$, which contains n_q de-meaned quarterly and n_m de-meaned monthly variables (i.e. $n = n_q + n_m$), be defined as

$$\tilde{\mathbf{y}}_t = \begin{bmatrix} y^q_{1,t} \\ \vdots \\ y^q_{n_q,t} \\ y^m_{1,t} - \rho^m_{1,1} y^m_{1,t-1} - \rho^m_{1,2} y^m_{1,t-2} \\ \vdots \\ y^m_{n_m,t} - \rho^m_{n_m,1} y^m_{n_m,t-1} - \rho^m_{n_m,2} y^m_{n_m,t-2} \end{bmatrix},$$

so that the system is written out in terms of the *quasi-differences* of the monthly indicators. Given this re-defined vector of observables, we cast our model into the following state space form:

$$\begin{aligned} \tilde{\mathbf{y}}_\mathbf{t} &= \mathbf{H}\mathbf{X}_t + \tilde{\boldsymbol{\eta}}_\mathbf{t}, & \tilde{\boldsymbol{\eta}}_\mathbf{t} &\sim N(0, \tilde{\mathbf{R}}_\mathbf{t}) \\ \mathbf{X}_t &= \mathbf{F}\mathbf{X}_{t-1} + \mathbf{e}_t, & \mathbf{e}_t &\sim N(0, \mathbf{Q}_t) \end{aligned}$$

where the state vector is defined as $\mathbf{X}'_t = [a_t, \ldots, a_{t-4}, f_t, \ldots, f_{t-4}, \mathbf{u}^{q\prime}_t, \ldots, \mathbf{u}^{q\prime}_{t-4}]$. Setting $\lambda_1 = 1$ for identification, the matrices of parameters $\mathbf{H}$ and $\mathbf{F}$, are then constructed as shown below:

$$\mathbf{H} = \begin{bmatrix} [ccc|ccc|ccc] & \mathbf{H}_a & \begin{matrix} \mathbf{H}_{\lambda_q} \\ \mathbf{H}_{\lambda_m} \end{matrix} & \mathbf{H}_u \end{bmatrix},$$

where the respective blocks of $\mathbf{H}$ are defined as

$$\mathbf{H}_a = \left[\begin{array}{ccccc} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ \hline \multicolumn{5}{c}{\mathbf{0}_{(n-1)\times 5}} \end{array}\right], \quad \mathbf{H}_{\lambda_q} = \begin{bmatrix} 1 & \lambda_2 & \ldots & \lambda_{n_q} \end{bmatrix}' \times \begin{bmatrix} 1/3 & 2/3 & 1 & 2/3 & 1/3 \end{bmatrix},$$

$$\mathbf{H}_{\lambda_m} = \begin{bmatrix} [c|ccc]\lambda_{n_q+1} - \lambda_{n_q+1}\rho^m_{1,1} - \lambda_{n_q+1}\rho^m_{1,2} & \mathbf{0}_{1\times 4} \\ \vdots & \vdots \\ \lambda_n - \lambda_n \rho^m_{n_m,1} - \lambda_n \rho^m_{n_m,2} & \mathbf{0}_{1\times 4} \end{bmatrix}$$

$$\mathbf{H}_u = \begin{bmatrix} \bar{\mathbf{H}}_u \\ \mathbf{0}_{n_m \times 5} \end{bmatrix}, \qquad \bar{\mathbf{H}}_u = \mathbf{1}_{n_q \times 1} \times \begin{bmatrix} 1/3 & 2/3 & 1 & 2/3 & 1/3 \end{bmatrix},$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0} & & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_2 & & & \\ \vdots & & \mathbf{F}_{2+1} & & \vdots \\ \vdots & & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \mathbf{0} & \mathbf{F}_{2+n_q} \end{bmatrix},$$

where the respective blocks of $\mathbf{F}$ are defined as

$$\mathbf{F}_1 = \left[\begin{array}{cc} 1 & \mathbf{0}_{1\times 4} \\ \hline \mathbf{I}_4 & \mathbf{0}_{4\times 1} \end{array}\right] \qquad \mathbf{F}_2 = \left[\begin{array}{ccc} \phi_1 & \phi_2 & \mathbf{0}_{1\times 3} \\ \hline \multicolumn{2}{c}{\mathbf{I}_4} & \mathbf{0}_{4\times 1} \end{array}\right] \qquad \mathbf{F}_{2+j} = \left[\begin{array}{ccc} \rho^q_{j,1} & \rho^q_{j,2} & \mathbf{0}_{1\times 3} \\ \hline \multicolumn{2}{c}{\mathbf{I}_4} & \mathbf{0}_{4\times 1} \end{array}\right]$$

for $j = 1, ..., n_q$.

The error terms are denoted as

$$\begin{aligned} \tilde{\boldsymbol{\eta}}_t &= [\mathbf{0}_{1\times n_q}, \tilde{\boldsymbol{\eta}}_t^{\boldsymbol{m}\prime}]' \\ \mathbf{e}_t &= \begin{bmatrix} v_{a_t} & \mathbf{0}_{4\times 1} & \epsilon_t & \mathbf{0}_{4\times 1} & \eta_{1,t} & \mathbf{0}_{4\times 1} & \dots & \eta_{n_q,t} & \mathbf{0}_{4\times 1} \end{bmatrix}', \end{aligned}$$

with covariance matrices

$$\tilde{\mathbf{R}}_{\mathbf{t}} = \begin{bmatrix} \mathbf{0}_{n_q \times n_q} & \mathbf{0}_{n_q \times n_m} \\ \mathbf{0}_{n_m \times n_q} & \mathbf{R}_t \end{bmatrix},$$

where $\mathbf{R}_t = diag(\sigma^2_{\eta^m_{1,t}}, \dots, \sigma^2_{\eta^m_{n_m,t}})$ and

$$\mathbf{Q}_t = diag(\omega_a^2, \mathbf{0}_{1\times 4}, \sigma^2_{\epsilon,t}, \mathbf{0}_{1\times 4}, \sigma^2_{\eta^q_{1,t}}, \mathbf{0}_{1\times 4}, \dots, \sigma^2_{\eta^q_{n_q,t}}, \mathbf{0}_{1\times 4}).$$

Details of the Gibbs Sampler

For ease of notation, we restrict the description to the case $m = r = 1$, $\mathbf{B} = 1$ and $\mathbf{a}_t = a_t$. Let $\boldsymbol{\theta} \equiv \{\boldsymbol{\lambda}, \boldsymbol{\Phi}, \boldsymbol{\rho}, \omega_a, \omega_\varepsilon, \omega_{\eta_1}, \ldots, \omega_{\eta_n}\}$ be a vector that collects the underlying parameters, where $\boldsymbol{\Phi}$ and $\boldsymbol{\rho}$ contain the parameters for factor and idiosyncratic components respectively. The model is estimated using a Markov Chain Monte Carlo (MCMC) Gibbs sampling algorithm in which conditional draws of the latent variables, $\{a_t, f_t\}_{t=1}^T$, the parameters, $\boldsymbol{\theta}$, and the stochastic volatilities, $\{\sigma_{\varepsilon,t}, \sigma_{\eta_{i,t}}\}_{t=1}^T$ are obtained sequentially. The algorithm has a block structure composed of the following steps.

0. Initialization
The model parameters are initialized at arbitrary starting values $\boldsymbol{\theta}^0$, and so are the sequences for the stochastic volatilities, $\{\sigma_{\varepsilon,t}^0, \sigma_{\eta_{i,t}}^0\}_{t=1}^T$. Set $j = 1$.

1. Draw latent variables conditional on model parameters and SVs
Obtain a draw $\{a_t^j, f_t^j, \mathbf{u}_t^q\}_{t=1}^T$ from $p(\{a_t, f_t\}_{t=1}^T | \boldsymbol{\theta}^{j-1}, \{\sigma_{\varepsilon,t}^{j-1}, \sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T, \mathbf{y})$.

This step of the algorithm uses the state space representation described above (Appendix 4.7.4), and produces a draw from the entire state vector $\mathbf{X}_t$ (which includes the long-run growth components, a_t, the common factor, f_t, and the idiosyncratic components of the quarterly variables, $\mathbf{u}_t^q$) by means of a forward-filtering backward-smoothing algorithm, see Carter and Kohn (1994) or Kim and Nelson (1999b). In particular, we adapt the algorithm proposed by Bai and Wang (2015), which is robust to numerical inaccuracies, and extend it to the case with mixed frequencies and missing data following Mariano and Murasawa (2003), as explained in section 4.3.3. Like Bai and Wang (2015), we initialise the Kalman Filter step from a normal distribution whose moments are independent of the model parameters, in particular $\mathbf{X}_0 \sim N(0, 10^4\mathbf{I})$.

2. Draw the variance of the time-varying GDP growth component
Obtain a draw $\omega_a^{2,j}$ from $p(\omega_a^2 | \{a_t^j\}_{t=1}^T)$.

Taking the sample $\{a_t^j\}_{t=1}^T$ drawn in the previous step as given, and posing an inverse-gamma prior $p(\omega_a^2) \sim IG(S_a, v_a)$ the conditional posterior of ω_a^2 is also drawn inverse-gamma distribution. As discussed in Section 4.4.2, we choose the scale $S_a = 10^{-3}$ and degrees of freedom $v_a = 1$ for our baseline specification.

3. Draw the autoregressive parameters of the factor VAR
Obtain a draw $\boldsymbol{\Phi}^j$ from $p(\boldsymbol{\Phi} | \{f_t^{j-1}, \sigma_{\varepsilon,t}^{j-1}\}_{t=1}^T)$.

Taking the sequences of the common factor $\{f_t^{j-1}\}_{t=1}^T$ and its stochastic volatility $\{\sigma_{\varepsilon,t}^{j-1}\}_{t=1}^T$ from previous steps as given, and posing a non-informative prior, the corresponding conditional posterior is drawn from the Normal distribution, see, e.g. Kim and Nelson (1999b). In the more general case of more than one factor, this step would be equivalent to drawing from the coefficients of a Bayesian VAR. Like Kim and Nelson (1999b), or Cogley and Sargent (2005), we reject draws which imply autoregressive coefficients in the explosive region.

4. Draw the factor loadings
Obtain a draw of $\boldsymbol{\lambda}^j$ from $p(\boldsymbol{\lambda}|\rho^{j-1}, \{f_t^{j-1}, \sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T, \mathbf{y})$.

Conditional on the draw of the common factor $\{f_t^{j-1}\}_{t=1}^T$, the measurement equations reduce to n independent linear regressions with heteroskedastic and serially correlated residuals. By conditioning on $\boldsymbol{\rho}^{j-1}$ and $\sigma_{\eta_{i,t}}^{j-1}$, the loadings can be estimated using GLS and non-informative priors. When necessary, we apply restrictions on the loadings using the formulas provided by de Wind and Gambetti (2014), see Appendix 4.7.5 for further information.

5. Draw the serial correlation coefficients of the idiosyncratic components
Obtain a draw of $\boldsymbol{\rho}^j$ from $p(\boldsymbol{\rho}|\lambda^{j-1}, \{f_t^{j-1}, \sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T, \mathbf{y})$.

Taking the sequence of the common factor $\{f_t^{j-1}\}_{t=1}^T$ and the loadings drawn in previous steps as given, the idiosyncratic components for the monthly variables can be obtained as $u_{i,t} = y_{i,t} - \lambda^{j-1} f_t^{j-1}$. For the quarterly variables, a draw of the idiosyncratic components has been obtained directly from Step 1. Given a sequence for the stochastic volatility of the i^{th} component, $\{\sigma_{\eta_{i,t}}^{j-1}\}_{t=1}^T$, the residual is standardized to obtain an autoregression with homoskedastic residuals whose conditional posterior can be drawn from the Normal distribution as in step 2.3.

6. Draw the stochastic volatilities
Obtain a draw of $\{\sigma_{\varepsilon,t}^j\}_{t=1}^T$ and $\{\sigma_{\eta_{i,t}}^j\}_{t=1}^T$ from $p(\{\sigma_{\varepsilon,t}\}_{t=1}^T|\boldsymbol{\Phi}^{j-1}, \{f_t^{j-1}\}_{t=1}^T)$, and from $p(\{\sigma_{\eta_{i,t}}\}_{t=1}^T|\boldsymbol{\lambda}^{j-1}, \boldsymbol{\rho}^{j-1}, \{f_t^{j-1}\}_{t=1}^T, \mathbf{y})$ respectively.

Finally, we draw the stochastic volatilities of the innovations to the factor and the idiosyncratic components independently, using the algorithm proposed by Kim et al. (1998), which uses a mixture of normal random variables to approximate the elements of the log-variance. This is a more efficient alternative to the exact Metropolis-Hastings algorithm previously proposed by Jacquier et al. (2002). For the general case in which there is more than one factor, the volatilities of the factor VAR can be drawn jointly, see Primiceri (2005).

Increase j by 1, go to Step 2.1 and iterate until convergence is achieved.

4.7.5 Implementing linear restrictions on the factor loadings

To impose linear restrictions on the factor loadings $\boldsymbol{\lambda}$ in equation (4.1) of the paper, we follow de Wind and Gambetti (2014). For linear restrictions of the form

$$\mathbf{R}\boldsymbol{\lambda} = \mathbf{r} \tag{4.11}$$

these authors consider the special case with $\mathbf{r} = \mathbf{0}$ in equation (54) in the appendix to their paper. For $\mathbf{r} \neq \mathbf{0}$, this equation is amended as shown here. Let $\boldsymbol{\lambda}^u$ and $\boldsymbol{\lambda}^r$ denote the unrestricted and restricted loading matrix, respectively. $\boldsymbol{\lambda}^r$ is then drawn from a posterior distribution defined by (4.12) to (4.14):

$$\boldsymbol{\lambda}^r \sim N\left(\overline{\boldsymbol{\lambda}}^r, \mathbf{P}_\lambda^r\right), \tag{4.12}$$

where

$$\begin{aligned} \overline{\boldsymbol{\lambda}}^r &= \boldsymbol{\lambda}^u - \mathbf{P}_\lambda^u \mathbf{R}' \left(\mathbf{R}\mathbf{P}_\lambda^u \mathbf{R}'\right)^{-1} \left(\mathbf{R}\boldsymbol{\lambda}^u - \mathbf{r}\right) & (4.13) \\ \mathbf{P}_\lambda^r &= \mathbf{P}_\lambda^u - \mathbf{P}_\lambda^u \mathbf{R}' \left(\mathbf{R}\mathbf{P}_\lambda^u \mathbf{R}'\right)^{-1} \mathbf{R}\mathbf{P}_\lambda^u. & (4.14) \end{aligned}$$

4.7.6 Details on the Construction of the Data Base

US (Vintage) Data Base

For our US real-time forecasting evaluation, we consider data vintages since 11 January 2000 capturing the real activity variables listed in the text. For each vintage, the start of the sample is set to January 1960, appending missing observations to any series which starts after that date. All times series are obtained from one of these sources: (1) Archival Federal Reserve Economic Data (ALFRED), (2) Bloomberg, (3) Haver Analytics. Table 4.5 provides details on each series, including the variable code corresponding to the different sources.

For several series, in particular Retail Sales, New Orders, Imports and Exports, only vintages in nominal terms are available, but series for appropriate deflators are available from Haver, and these are not subject to revisions. We therefore deflate them using, respectively, CPI, PPI for Capital Equipment, and Imports and Exports price indices. Additionally, in several occasions the series for New Orders, Personal Consumption, Vehicle Sales and Retail Sales are subject to methodological changes and part of their history gets discontinued. In this case, given our interest in using long samples for all series, we use older vintages to splice the growth rates back to the earliest possible date.

For *soft* variables real-time data is not as readily available. The literature on real-time forecasting has generally assumed that these series are unrevised, and therefore used the latest available vintage. However while the underlying survey responses are indeed not revised, the seasonal adjustment procedures applied to them do lead to important differences between the series as was available at the time and the latest vintage. For this reason we use seasonally un-adjusted data and re-apply the Census-X12 procedure in real time to obtain a real-time seasonally adjusted version of the surveys. We follow the same procedure for the initial unemployment claims series. We then use Bloomberg to obtain the exact date in which each monthly data point was first published.

Table 4.5: DETAILED DESCRIPTION OF US DATA SERIES

	Frequ.	Start Date	Vintage Start	Trans- formation	Publ. Lag	Data Code
Real Gross Domestic Product	Q	Q2:1947	Dec 91	%QoQ Ann	26	GDPC1(F)
Real Consumption (ex. durables)	Q	Q2:1947	Dec 91	%QoQ Ann	26	
Hours worked	Q	Q2:1948	Dec 91	%QoQ Ann	28	
Real Investment (incl. durable cons.)	Q	Q2:1947	Dec 91	%QoQ Ann	26	
Real Industrial Production	M	Jan 47	Jan 97	% MoM	15	INDPRO(F)
Real Manufacturers' New Orders Nondefense Capital Goods Excluding Aircraft	M	Mar 68	Mar 97	% MoM	25	NEWORDER(F)[1] PPICPE(F)
Real Light Weight Vehicle Sales	M	Feb 67	Mar 97	% MoM	1	ALTSALES(F)[2] TLVAR(H)
Real Personal Income less Transfer Payments	M	Feb 59	Dec 97	% MoM	27	DSPIC96(F)
Real Retail Sales Food Services	M	Feb 47	Jun 01	% MoM	15	RETAIL(F) CPIAUCSL(F) RRSFS(F)[3]
Real Exports of Goods	M	Feb 68	Jan 97	% MoM	35	BOPGEXP(F)[4] C111CPX(H) TMXA(H)
Real Imports of Goods	M	Feb 69	Jan 97	% MoM	35	BOPGIMP(F)[4] C111CP(H) TMMCA(H)
Building Permits	M	Feb 60	Aug 99	% MoM	19	PERMIT(F)
Housing Starts	M	Feb 59	Jul 70	% MoM	26	HOUST(F)
New Home Sales	M	Feb 63	Jul 99	% MoM	26	HSN1F(F)
Total Nonfarm Payroll Employment (Establishment Survey)	M	Jan 47	May 55	% MoM	5	PAYEMS(F)
Civilian Employment (Household Survey)	M	Feb 48	Feb 61	% MoM	5	CE16OV(F)
Unemployed	M	Feb 48	Feb 61	% MoM	5	UNEMPLOY(F)
Initial Claims for UE	M	Feb 48	Jan 00*	% MoM	4	LICM(H)

(Continues on next page)

Detailed Description of US Data Series (Continued)

Markit Manufacturing PMI	M	May 07	Jan 00*	-	-7	S111VPMM(H)[5] H111VPMM(H)
ISM Manufacturing PMI	M	Jan 48	Jan 00*	-	1	NMFBAI(H) NMFNI(H) NMFEI(H) NMFVDI(H)[6]
ISM Non-manufacturing PMI	M	Jul 97	Jan 00*	-	3	NAPMCN(H)
Conference Board: Consumer Confidence	M	Feb 68	Jan 00*	Diff 12 M.	-5	CCIN(H)
University of Michigan: Consumer Sentiment	M	May 60	Jan 00*	Diff 12 M.	-15	CSENT(H)[5] CONSSENT(F) Index(B)
Richmond Fed Manufacturing Survey	M	Nov 93	Jan 00*	-	-5	RIMSXN(H) RIMNXN(H) RIMLXN(H)[6]
Philadelphia Fed Business Outlook	M	May 68	Jan 00*	-	0	BOCNOIN(H) BOCNONN(H) BOCSHNN(H) BOCDTIN(H) BOCNENN(H)[6]
Chicago PMI	M	Feb 67	Jan 00*	-	0	PMCXPD(H) PMCXNO(H) PMCXI(H) PMCXVD(H)[6]
NFIB: Small Business Optimism Index	M	Oct 75	Jan 00*	Diff 12 M.	15	NFIBBN (H)
Empire State Manufacturing Survey	M	Jul 01	Jan 00*	-	-15	EMNHN(H) EMSHN(H) EMDHN(H) EMDSN(H) EMESN(H)[6]

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. In the last column, (B) = Bloomberg; (F) = FRED; (H) = Haver;
1) deflated using PPI for capital equipment; 2) for historical data not available in ALFRED we used data coming from HAVER; 3) using deflated nominal series up to May 2001 and real series afterwards; 4) nominal series from ALFRED and price indices from HAVER. For historical data not available in ALFRED we used data coming from HAVER; 5) preliminary series considered; 6) NSA subcomponents needed to compute the SA headline index. * Denotes seasonally un-adjusted series which have been seasonally adjusted in real time.

Data Base for Other G7 Economies

Table 4.6: Data used for Canada

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Industrial Production: Manuf., Mining, Util.	M	Jan-1960	% MoM
Manufacturing New Orders	M	Feb-1960	% MoM
Manufacturing Turnover	M	Feb-1960	% MoM
New Passenger Car Sales	M	Jan-1960	% MoM
Real Retail Sales	M	Feb-1970	% MoM
Construction: Dwellings Started	M	Feb-1960	% MoM
Residential Building Permits Auth.	M	Jan-1960	% MoM
Real Exports	M	Jan-1960	% MoM
Real Imports	M	Jan-1960	% MoM
Unemployment Ins.: Initial and Renewal Claims	M	Jan-1960	% MoM
Employment: Industrial Aggr. excl. Unclassified	M	Feb-1991	% MoM
Employment: Both Sexes, 15 Years and Over	M	Feb-1960	% MoM
Unemployment: Both Sexes, 15 Years and Over	M	Feb-1960	% MoM
Consumer Confidence Indicator	M	Jan-1981	Diff 12 M.
Ivey Purchasing Managers Index	M	Jan-2001	Level
ISM Manufacturing PMI	M	Jan-1960	Level
University of Michigan: Consumer Sentiment	M	May-1960	Diff 12 M.

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table 4.7: DATA USED FOR GERMANY

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Mfg Survey: Production: Future Tendency	M	Jan-1960	Level
Ifo Demand vs. Prev. Month: Manufact.	M	Jan-1961	Level
Ifo Business Expectations: All Sectors	M	Jan-1991	Level
Markit Manufacturing PMI	M	Apr-1996	Level
Markit Services PMI	M	Jun-1997	Level
Industrial Production	M	Jan-1960	% MoM
Manufacturing Turnover	M	Feb-1960	% MoM
Manufacturing Orders	M	Jan-1960	% MoM
New Truck Registrations	M	Feb-1963	% MoM
Total Unemployed	M	Feb-1962	% MoM
Total Domestic Employment	M	Feb-1981	% MoM
Job Vacancies	M	Feb-1960	% MoM
Retail Sales Volume excluding Motor Vehicles	M	Jan-1960	% MoM
Wholesale Vol. excl. Motor Veh. and Motorcycles	M	Feb-1994	% MoM
Real Exports of Goods	M	Feb-1970	% MoM
Real Imports of Goods	M	Feb-1970	% MoM

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table 4.8: DATA USED FOR JAPAN

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
TANKAN All Industries: Actual Business Cond.	Q	Sep-1974	Diff 1 M.
Markit Manufacturing PMI	M	Oct-2001	Level
Small Business Sales Forecast	M	Dec-1974	Level
Small/Medium Business Survey	M	Apr-1976	Level
Consumer Confidence Index	M	Mar-1973	Level
Inventory to Sales Ratio	M	Jan-1978	Level
Industrial Production: Mining and Manufact.	M	Jan-1960	% MoM
Electric Power Consumed by Large Users	M	Feb-1960	% MoM
New Motor Vehicle Registration: Trucks, Total	M	Feb-1965	Diff 1 M.
New Motor Vehicle Reg: Passenger Cars	M	May-1968	% MoM
Real Retail Sales	M	Feb-1960	% MoM
Real Department Store Sales	M	Feb-1970	% MoM
Real Wholesale Sales: Total	M	Aug-1978	% MoM
Tertiary Industry Activity Index	M	Feb-1988	% MoM
Labor Force Survey: Total Unemployed	M	Jan-1960	% MoM
Overtime Hours / Total Hours (manufact.)	M	Feb-1990	% MoM
New Job Offers excl. New Graduates	M	Feb-1963	% MoM
Ratio of New Job Openings to Applications	M	Feb-1963	% MoM
Ratio of Active Job Openings and Active Job Appl.	M	Feb-1963	% MoM
Building Starts, Floor Area: Total	M	Feb-1965	% MoM
Housing Starts: New Construction	M	Feb-1960	% MoM
Real Exports	M	Feb-1960	% MoM
Real Imports	M	Feb-1960	% MoM

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table 4.9: Data used for the United Kingdom

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Mar-1960	% QoQ Ann.
Dist. Trades: Total Vol. of Sales	M	Jul-1983	Level
Dist. Trades: Retail Vol. of Sales	M	Jul-1983	Leve
CBI Industrial Trends: Vol. of Output Next 3 M.	M	Feb-1975	Level
BoE Agents' Survey: Cons. Services Turnover	M	Jul-1997	Level
Markit Manufacturing PMI	M	Jan-1992	Level
Markit Services PMI	M	Jul-1996	Level
Markit Construction PMI	M	Apr-1997	Level
GfK Consumer Confidence Barometer	M	Jan-1975	Diff 12 M.
Industrial Production: Manufacturing	M	Jan-1960	% MoM
Passenger Car Registrations	M	Jan-1960	% MoM
Retail Sales Volume: All Retail incl. Autom. Fuel	M	Jan-1960	% MoM
Index of Services: Total Service Industries	M	Feb-1997	% MoM
Registered Unemployment	M	Feb-1960	% MoM
Job Vacancies	M	Feb-1960	% MoM
LFS: Unemployed: Aged 16 and Over	M	Mar-1971	% MoM
LFS: Employment: Aged 16 and Over	M	Mar-1971	% MoM
Mortgage Loans Approved: All Lenders	M	May-1993	% MoM
Real Exports	M	Feb-1961	% MoM
Real Imports	M	Feb-1961	% MoM

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table 4.10: Data used for France

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Industrial Production	M	Feb-1960	% MoM
Total Commercial Vehicle Registrations	M	Feb-1975	% MoM
Household Consumption Exp.: Durable Goods	M	Feb-1980	% MoM
Real Retail Sales	M	Feb-1975	% MoM
Passenger Cars	M	Feb-1960	% MoM
Job Vacancies	M	Feb-1989	% MoM
Registered Unemployment	M	Feb-1960	% MoM
Housing Permits	M	Feb-1960	% MoM
Housing Starts	M	Feb-1974	% MoM
Volume of Imports	M	Jan-1960	% MoM
Volume of Exports	M	Jan-1960	% MoM
Business Survey: Personal Prod. Expect.	M	Jun-1962	Level
Business Survey: Recent Output Changes	M	Jan-1966	Level
Household Survey: Household Conf. Indicator	M	Oct-1973	Diff 12 M.
BdF Bus. Survey: Production vs. Last M., Ind.	M	Jan-1976	Level
BdF Bus. Survey: Production Forecast, Ind.	M	Jan-1976	Level
BdF Bus. Survey: Total Orders vs. Last M., Ind.	M	Jan-1981	Level
BdF Bus. Survey: Activity vs. Last M., Services	M	Oct-2002	Level
BdF Bus. Survey: Activity Forecast, Services	M	Oct-2002	Level
Markit Manufacturing PMI	M	Apr-1998	Level
Markit Services PMI	M	May-1998	Level

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

Table 4.11: Data used for Italy

	Freq.	Start Date	Transformation
Real Gross Domestic Product	Q	Jun-1960	% QoQ Ann.
Markit Manufacturing PMI	M	Jun-1997	Level
Markit Services PMI: Business Activity	M	Jan-1998	Level
Production Future Tendency	M	Jan-1962	Level
ISTAT Services Survey: Orders, Next 3 M-	M	Jan-2003	Level
ISTAT Retail Trade Confidence Indicator	M	Jan-1990	Level
Industrial Production	M	Jan-1960	% MoM
Real Exports	M	Jan-1960	% MoM
Real Imports	M	Jan-1960	% MoM
Real Retail Sales	M	Feb-1990	% MoM
Passenger Car Registrations	M	Jan-1960	% MoM
Employed	M	Feb-2004	% MoM
Unemployed	M	Feb-1983	% MoM

Note: The second column refers to the sampling frequency of the data, which can be quarterly (Q) or monthly (M). % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. All series were obtained from the Haver Analytics database.

4.7.7 Choice of Priors

As explained in the paper, we use non-informative priors for loadings and serial correlation coefficients of factor and idiosyncratic components in order to aide comparability with the previous literature, which has generally used classical estimation methods. With respect to the choice of priors on the new parameters of our specification, namely ω_a^2, ω_ε^2 and $\omega_{\eta,i}^2$ in equations (4.5)-(4.7), we closely follow the related literature, in particular Cogley and Sargent (2005) and Primiceri (2005), by setting relatively conservative priors, which shrink the model towards the benchmark with no time-variation, but are still loose enough for the data to be able to speak. In particular, in all the inverse-gamma (IG) distributions we set the number of degrees of freedom to 1, the minimum required to make the prior distributions proper while keeping the weight of the prior low. As to the choice of the scale parameter of the IG distributions, it is worth pointing out that this does not parametrize time variation itself, but rather incorporates a prior belief about the amount of time variation. To gain an intuition about the prior on ω_a^2, in Section 4.4.2 we note that the chosen value of 0.001 implies that over a period of ten years the random walk process of the long-run growth rate is expected to vary with a standard deviation of around 0.4 percentage points in annualized terms, which we believe is a fairly conservative prior in terms of economic magnitudes. The choice of 10^{-4} for the prior on ω_ε^2 and $\omega_{\eta,i}^2$ is similar to the approach of Primiceri (2005).

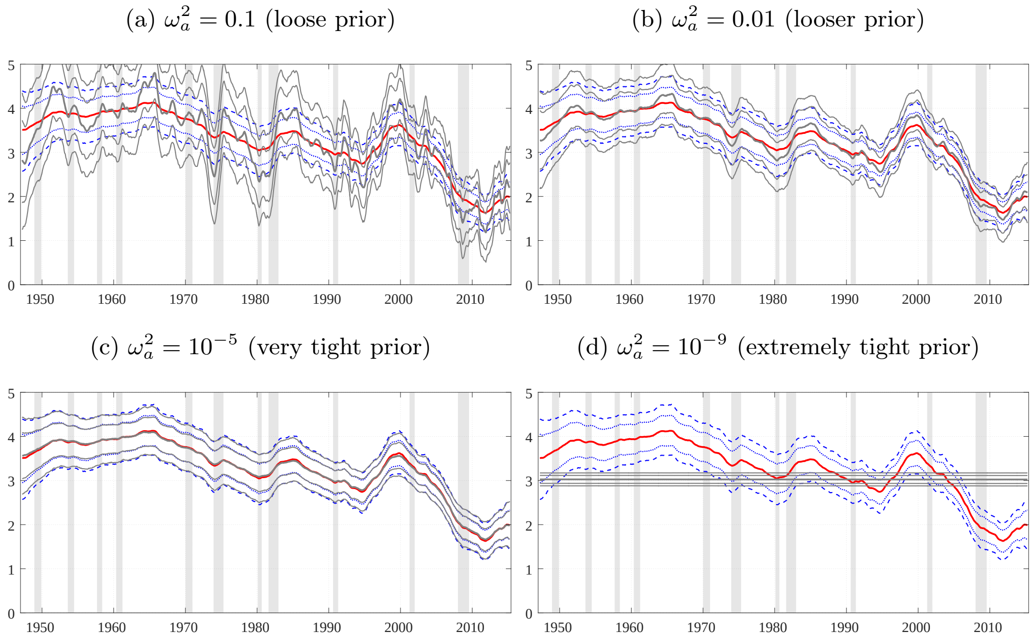
To shed some light on the robustness of our results to the choice of priors, in what follows we explore the sensitivity of our main results to varying the tightness of the respective priors. To summarize the most notable finding, we find that the data strongly drives the result of time variation both in the long-run growth rate and the volatilies: a dogmatically large amount of shrinkage is needed in order to make either of them disappear.

Robustness checks on prior choice

Prior on innovation variance to the time-varying long-run growth rate

In Figure 4.16 we explore the sensitivity of our key results to the choice of the scale parameter of the prior on the innovation variance to the time-varying long-run growth rate of real GDP, ω_a^2. Each panel plots our baseline estimate of g_t, which has been obtained with a prior scale of 10^{-3} (red/blue). We then successively superimpose alternative estimates obtained when imposing both looser and tighter prior scales (gray). Panel (a) of the figure reveals that with a prior implying a very large variance the estimated trend is pinned down with relatively more uncertainty and evolves in bumpy fashion, yet the qualitative pattern around the evolution of long-run growth, remains clearly visible. Panels (b) and (c) show that using a ten times looser prior (0.01) and a hundred times tighter prior (10^{-5}) than the one in our baseline setting gives very similar results to ours. In the later case, the estimate is almost identical. Finally, a dogmatically tight prior (10^{-9}) is required to make variation in the long-run growth rate disappear entirely, which is visible in Panel (d).

Figure 4.16: Comparison Across Different Prior Scales of ω_{α}^2



Note: In each panel our baseline the median estimate of real GDP growth based on a scale of 10^{-3} is presented (red), with corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals. The corresponding estimates based an different prior scales are superimposed in gray in each panel.

Prior on innovation variance to the SV

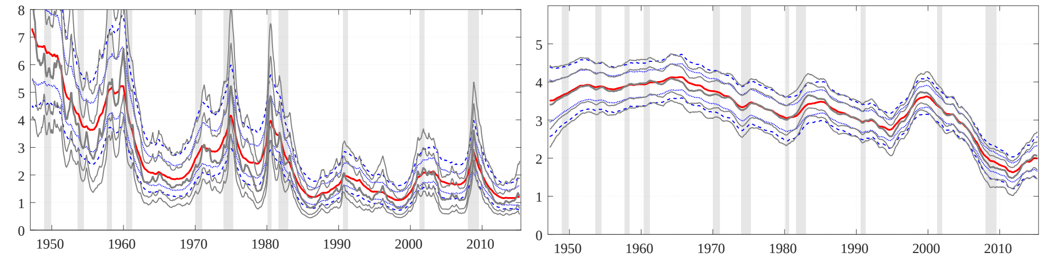
Figure 4.17 presents the sensitivity of the results to the choice of the scale parameter of the prior on the innovation variance to the SV in both the common factor and the idiosyncratic components. Similar to Figure 4.16 we compare our baseline estimates (red/blue), where we set $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 10^{-4}$, with estimates obtained under a range of varied prior scales (gray). In each case, the figure shows both the estimated SV of the factor as well as the estimate of the long-run growth rate of real GDP growth. Panel (a) displays the results for a very loose prior (1), while Panel (b) for a prior which is ten times looser than the baseline (10^{-3}). Finally, the estimates shown in Panel (c) are obtained under a tighter prior (10^{-5}). Again, the results reported in the paper do not seem to be affected. Both the estimates of the SV and the long-run growth rate of real GDP are almost identical to our main results.

Prior on serial correlation in factor and idiosyncratic components

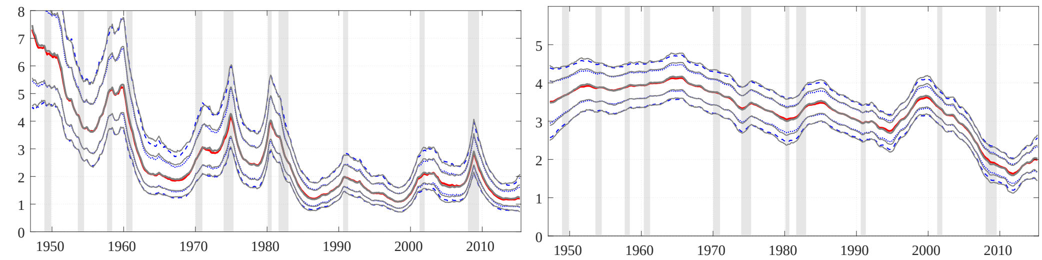
As a final robustness check, we consider "Minnesota"-style priors on the autoregressive coefficients of the factor as well as shrinking the coefficients of the serial correlation towards zero. To be precise, we center the prior on the first lag of the factor around 0.9 and all other lags at zero. The motivation for these priors is to express a preference for a more parsimonious model where the factors capture the bulk of the persistence of the series and the idiosyncratic components are close to iid, that is closer to true measurement error. These alternative priors do not mean-

Figure 4.17: Comparison Across Different Prior Scales of ω_{ε}^2 and $\omega_{\eta,i}^2$

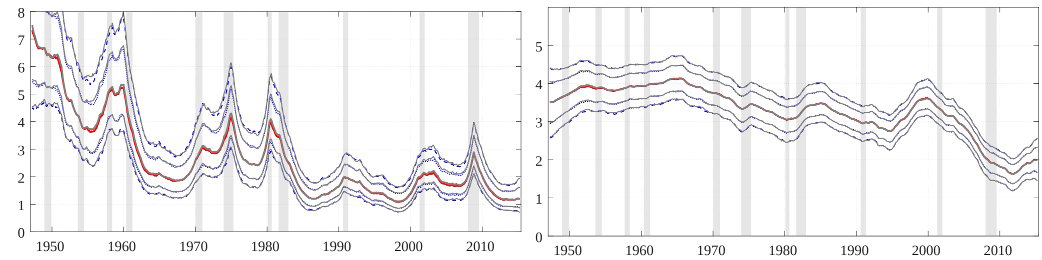
(a) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 1$ (very loose prior)



(b) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 10^{-3}$ (looser prior)



(c) $\omega_{\varepsilon}^2 = \omega_{\eta,i}^2 = 10^{-5}$ (tighter prior)



Note: In each panel our baseline estimate of the SV of the common factor based on a scale of 10^{-4} is presented (red) in the left chart. The right chart plots the estimate of the long-run growth rate of real GDP based on the same scale. Corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals are also plotted. The analogue estimates based on the alternative prior scales are superimposed in gray in each panel.

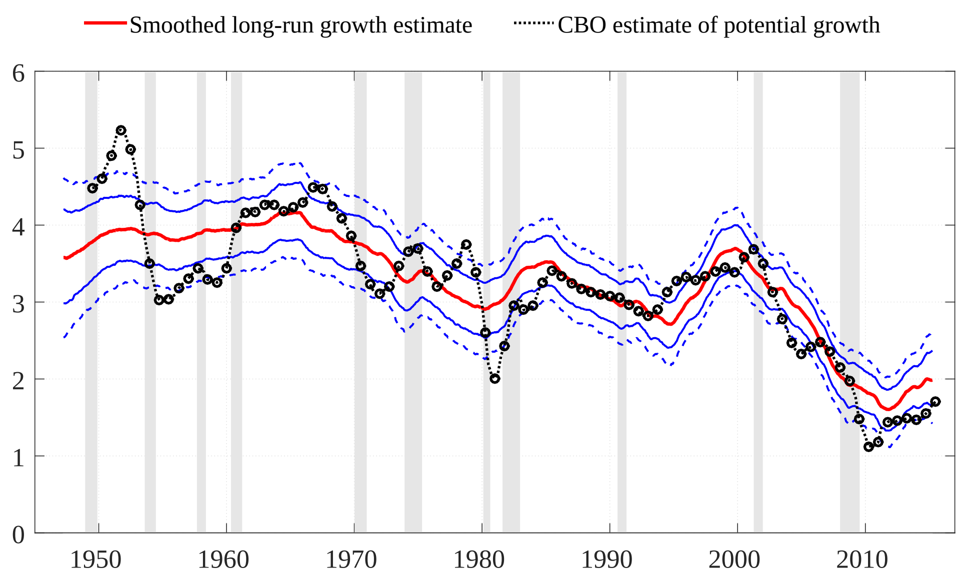
ingfully affect the posterior estimates of our main objects of interest, so we omit additional figures. Note that we have found some evidence that the use of such priors might at times improve the convergence of the algorithm. Specifically, when we apply the model to the other G7 economies (see Section 4.5), we find that for some countries where few monthly indicators are available, shrinking the serial correlations of the idiosyncratic components towards zero helps obtaining a common factor that is persistent.

4.7.8 Comparison with CBO Measure of Potential Output

Our estimate of long-run growth and the CBO's potential growth estimate capture related but not identical concepts. The CBO measures the growth rate of potential output, i.e. the level of output that could be obtained if all resources were used fully, whereas our estimate, similar to Beveridge and Nelson (1981), measures the component of the growth rate that is expected to be permanent. Moreover, the CBO estimate is constructed using the so-called "production function approach", which is radically different from the DFM methodology.[51]

As a simple sanity check, it is interesting to see that despite employing different statistical methods they produce qualitatively similar results, visible in Figure 4.18, with the CBO estimate displaying a more marked cyclical pattern but remaining for most of the sample within the 90% credible posterior interval of our estimate. As in our estimate, most of the slowdown occurred prior to the Great Recession. The CBO's estimate was below ours immediately after the recession, reaching an unprecedented low level of about 1.25% in 2010, and remains in the lower bound of our posterior estimate since then. Section 4.4.6 expands on the reason for this divergence and argues that this is likely to stem from the larger amount of information incorporated in the DFM. In fact, the CBO estimate of potential growth is noticeably more cyclical.

Figure 4.18: Long-run GDP Growth Estimate in Comparison to CBO



Note: The figure displays the posterior median estimate of long-run GDP growth with the corresponding credible intervals, as displayed in Figure 4.2 Panel (a) in the main body of the paper, in comparison with the CBO's measure of potential output growth, which is shown in black circles.

[51] Essentially, the production function approach calculates the trend components of the supply inputs to a neoclassical production function (the capital stock, total factor productivity, and the total amount of hours) using statistical filters and then aggregates them to obtain an estimate of the trend level of output. See CBO (2001).

4.7.9 Details About the Forecast Evaluation

Using our real-time database of US vintages, we re-estimate the following three models each day in which new data is released: a benchmark with constant long-run GDP growth and constant volatility (Model 0, similar to Banbura and Modugno (2014)), a version with constant long-run growth but with stochastic volatility (Model 1, similar to Marcellino et al. (2014)), and the baseline model put forward in the paper with both time-variation in the long-run growth of real GDP and SV (Model 2). Allowing for an intermediate benchmark with only SV allows us to evaluate how much of the improvement in the model can be attributed to the addition of the long-run variation in GDP as opposed to the SV. We evaluate the point and density forecast accuracy relative to the initial ("Advance") release of GDP, which is released between 25 and 30 days after the end of the reference quarter.[52]

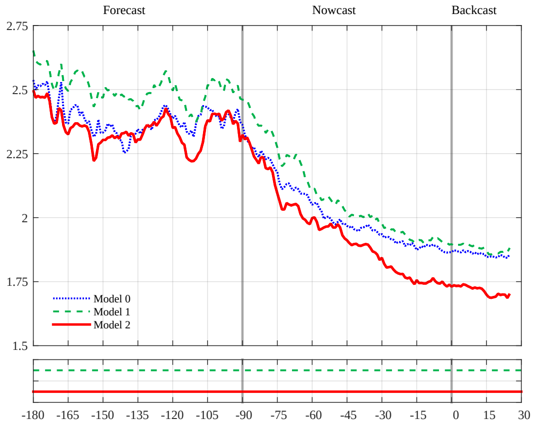
When comparing the three different models, we test the significance of any improvement of Models 1 and 2 relative to Model 0. This raises some important econometric complications given that (i) the three models are nested, (ii) the forecasts are produced using an expanding window, and (iii) the data used is subject to revision. These three issues imply that commonly used test statistics for forecasting accuracy, such as the one proposed by Diebold and Mariano (1995) and Giacomini and White (2006) will have a non-standard limiting distribution. However, rather than not reporting any test, we follow the "pragmatic approach" of Faust and Wright (2013) and Groen et al. (2013), who build on Monte Carlo results in Clark and McCracken (2012). Their results indicate that the Harvey et al. (1997) small sample correction of the Diebold and Mariano (1995) statistic results in a good sized test of the null hypothesis of equal finite sample forecast precision for both nested and non-nested models, including cases with expanded window-based model updating. Overall, the results of the tests should be interpreted more as a rough gauge of the significance of the improvement than a definitive answer to the question. We compute various point and density forecast accuracy measures at different moments in the release calendar, to assess how the arrival of information improves the performance of the model. In particular, the computations are carried out starting 180 days before the end of the reference quarter, and every subsequent day up to 25 days after its end, when the GDP figure for the quarter is usually released. This means that we will evaluate the forecasts of the next quarter, current quarter (nowcast), and the previous quarter (backcast). We consider two different samples for the evaluation: the full sample (2000:Q1-2015:Q1) and the sample covering the recovery since the Great Recession (2009:Q2-2015:Q1).

[52] We have explored the alternative of evaluating the forecasts against subsequent releases, or the latest available vintages. The relative performance of the three models is broadly unchanged, but all models do better at forecasting the initial release. If the objective is to improve the performance of the model relative to the first official release, then ideally an explicit model of the revision process would be desirable. The results are available upon request.

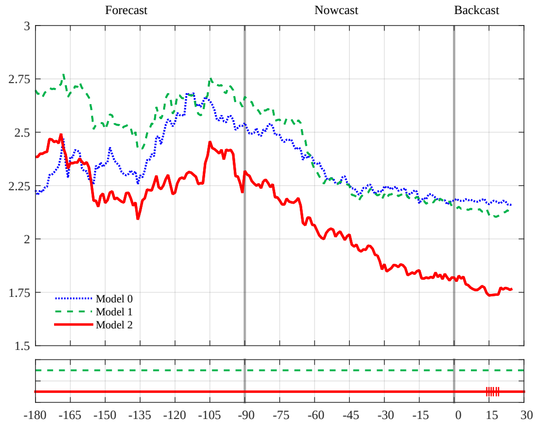
Figure 4.19: Point Forecast Accuracy Evaluation

(a) Root Mean Squared Error

Full Sample: 2000:Q1-2015:Q1

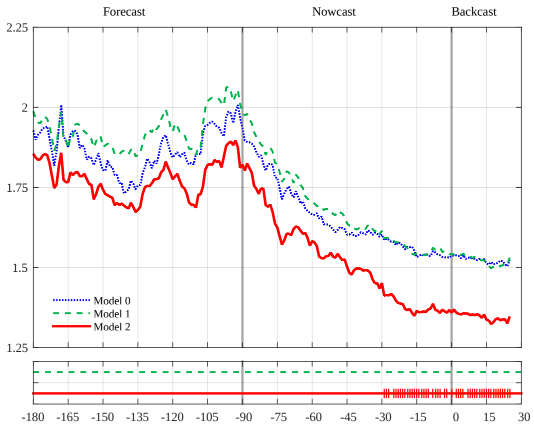


Recovery Sample: 2009:Q2-2015:Q1

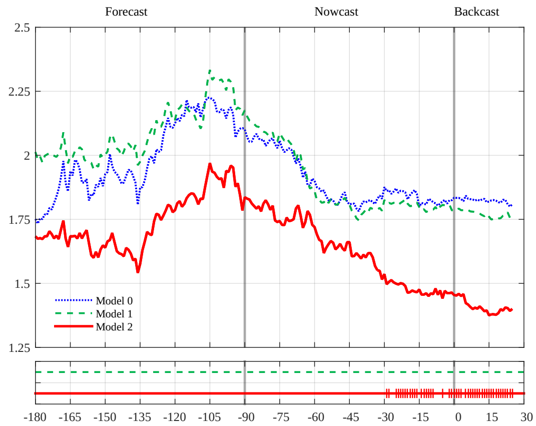


(b) Mean Absolute Error

Full Sample: 2000:Q1-2015:Q1



Recovery Sample: 2009:Q2-2015:Q1



Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

Point Forecast Evaluation

Figure 4.19 shows the results of evaluating the posterior mean as point forecast. We use two criteria, the root mean squared error (RMSE) and the mean absolute error (MAE). As expected, both of these decline as the quarters advance and more information on monthly indicators becomes available, see e.g. Banbura et al. (2012). Both the RMSE and the MAE of Model 2 are lower than that of Model 0, particularly so from the start of the nowcasting period, while Model 1 is somewhat worse overall. Our gauge of significance indicates that these differences in nowcasting performance are significant at the 10% level for the overall sample in the case of the MAE, but not the RMSE. The improvement in performance is much clearer in the recovery sample. In fact, the inclusion of the time varying long run component of GDP helps anchor GDP predictions at a level consistent with the weak recovery experienced in the past few years and produces nowcasts that are 'significantly' superior to those of the reference model from around 30 days before the end of the reference quarter. In essence, ignoring the variation in long-run GDP growth would have resulted in being on average around 1 percentage point too optimistic from 2009 to 2015.

Density Forecast Evaluation

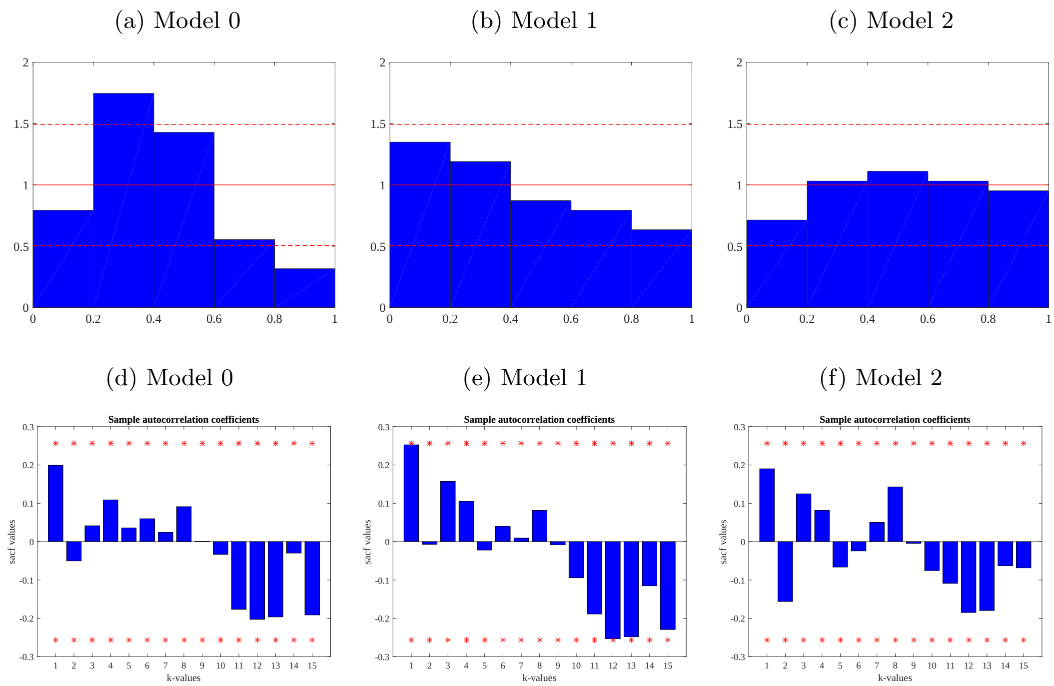
Density forecasts can be used to assess the ability of a model to predict unusual developments, such as the likelihood of a recession or a strong recovery given current information. The adoption of a Bayesian framework allows us to produce density forecasts from the DFM that consistently incorporate both filtering and estimation uncertainty. Figure 4.20 reports the probability integral transform (PITs) and the associated autocorrelation functions (ACFs) for the 3 models calculated with the nowcast of the last day of the quarter. Diebold et al. (1998) highlight that well calibrated densities are associated with uniformly distributed and independent PITs. Figure 4.20 suggests that the inclusion of SV is paramount to get well calibrated densities, whereas the inclusion of the long-run growth component helps to get a more appropriate representation of the right side of the distribution, as well as making sure that the first order autocorrelation is not statistically significant.

There are several measures available for density forecast evaluation. The (average) log score, i.e. the logarithm of the predictive density evaluated at the realization, is one of the most popular, rewarding the model that assigns the highest probability to the realized events. Gneiting and Raftery (2007), however, caution against using the log score, emphasizing that it does not appropriately reward values from the predictive density that are close but not equal to the realization, and that it is very sensitive to outliers. They therefore propose the use of the (average) continuous rank probability score (CRPS) in order to address these drawbacks of the log-score. Figure 4.21 shows that by both measures our model outperforms its counterparts. Interestingly, the comparison of Model 1 and Model 2 suggests that failing to properly account for the long-run growth component might give a misrepresentation of the GDP densities, resulting in poorer density forecasts.

In addition to the above results, we also assess how the three models fare when different areas of their predictive densities are emphasized in the forecast evaluation. To do that we follow Groen et al. (2013) and compute weighted averages of Gneiting and Raftery (2007) quantile scores (QS) that are based on quantile forecasts that correspond to the predictive densities from the different models (Figure 4.22).[53] Our results indicate that while there is an improvement in density nowcasting for the

[53]As Gneiting and Ranjan (2011) show, integrating QS over the quantile spectrum gives the CRPS.

Figure 4.20: Probability Integral Transform (PITs)



Note: The upper three panels display the cdf of the Probability Integral Transforms (PITs) evaluated on the last day of the reference quarter, while the lower three display the associated autocorrelation functions.

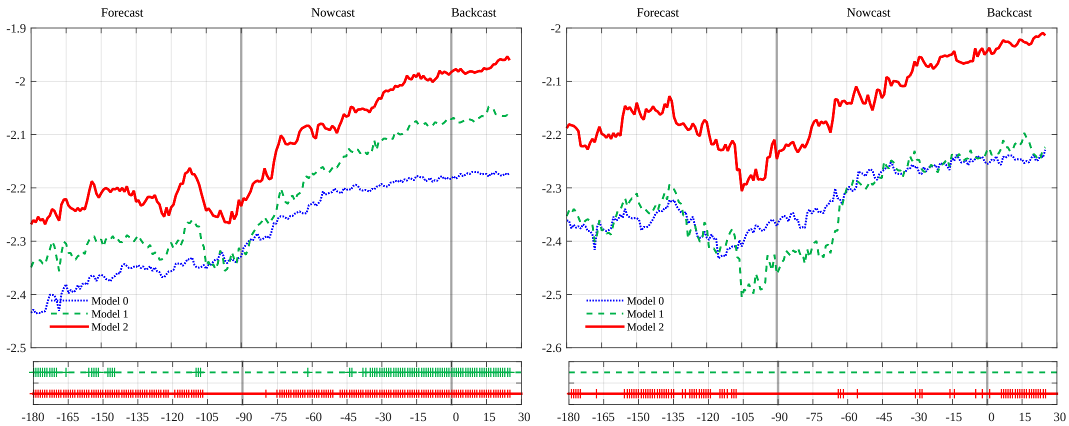
entire distribution, the largest improvement comes from the right tail. For the full sample, Model 1 is very close to Model 0, suggesting that being able to identify the location of the distribution is key to the improvement in performance. In order to appreciate the importance of the improvement in the density forecasts, and in particular in the right side of the distribution, we calculated a recursive estimate of the likelihood of a 'strong recovery', where this is defined as the probability of an average growth rate of GDP (over the present and next three quarters) above the historical average. Model 0 and Model 2 produce very similar probabilities up until 2011 when, thanks to the downward revision of long-run GDP growth, Model 2 starts to deliver lower probability estimates consistent with the observed weak recovery. The Brier score for Model 2 is 0.186 whereas the score for Model 0 is 0.2236 with the difference significantly different at 1% (Model 1 is essentially identical to Model 0).

In sum, the out-of-sample forecasting evaluation indicates that allowing for time-varying long-run GDP growth and SV produces short-run forecasts that are on average either similar to or improve upon the benchmark model. The performance tends to improve substantially in the sub-sample including the recovery from the Great Recession, coinciding with the significant downward revision of the model's assessment of long-run growth. The results indicate that while there is an improvement in density nowcasting for the entire distribution, the largest improvement comes from the right tail.

Figure 4.21: Density Forecast Accuracy Evaluation

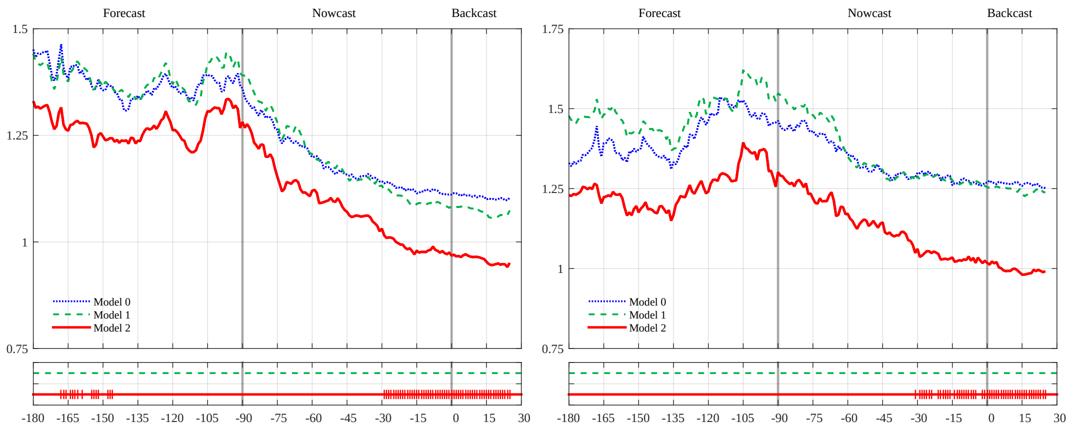
(a) Log Probability Score

Full Sample: 2000:Q1-2015:Q1 | Recovery Sample: 2009:Q2-2015:Q1



(b) Continuous Rank Probability Score

Full Sample: 2000:Q1-2015:Q1 | Recovery Sample: 2009:Q2-2015:Q1

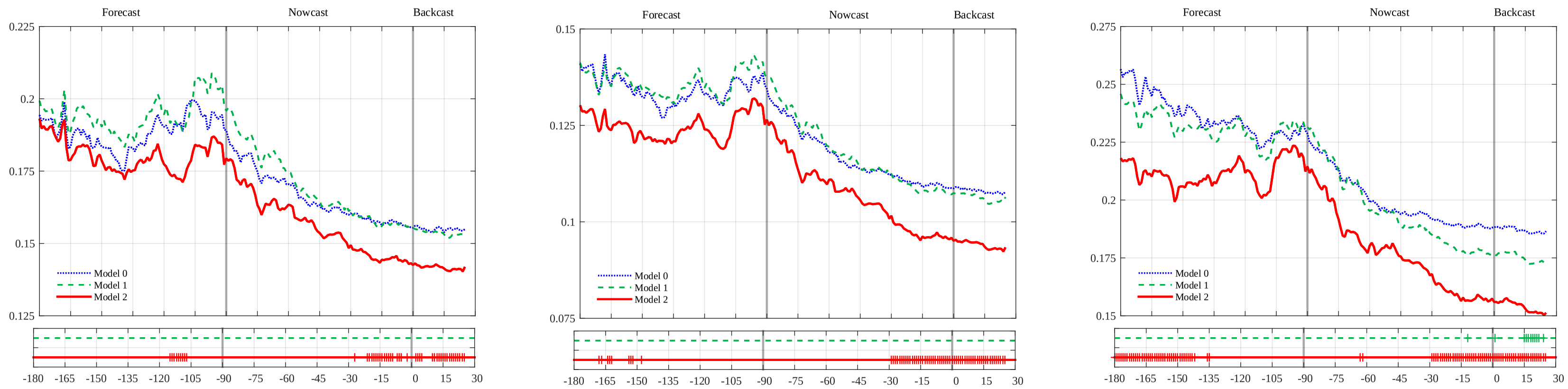


Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

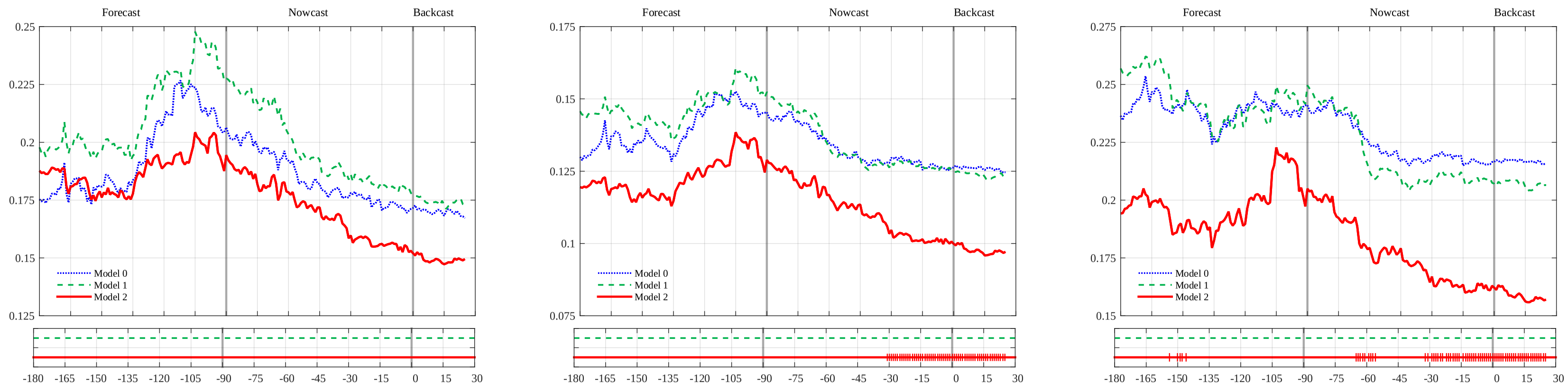
Figure 4.22: Density Forecast Accuracy Evaluation: Quantile Scores

Full Sample: 2000:Q1-2015:Q1

(a) Left (b) Center (c) Right



Recovery Sample: 2009:Q2-2015:Q1



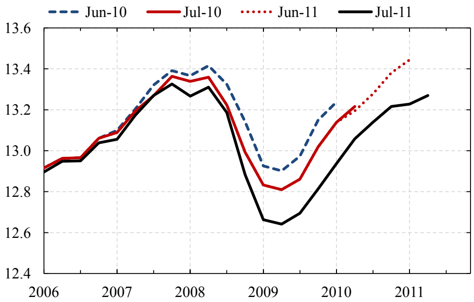
Note: The horizontal axis indicates the forecast horizon, expressed as the number of days to the end of the reference quarter. Thus, from the point of view of the forecaster, forecasts produced 180 to 90 days before the end of a given quarter are a forecast of next quarter; forecasts 90-0 days are nowcasts of current quarter, and the forecasts produced 0-25 days after the end of the quarter are backcasts of last quarter. The boxes below each panel display, with a vertical tick mark, a gauge of statistical significance at the 10% level of any difference with Model 0, for each forecast horizon, as explained in the main text.

4.7.10 Case Study - The Decline of The Long-Run Growth Estimate in Mid-2010 and Mid-2011

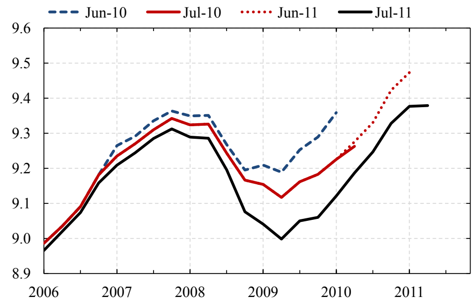
Figure 4.23 looks in more detail at the specific information that, in real time, led the model to reassess its estimate of long-run growth. There are large reassessments of long-run growth around July 2010 and July 2011, coinciding with the publication by the Bureau of Economic Analysis of the annual revisions to the National Accounts, which each year incorporate previously unavailable information for the previous three years. In both cases, the revisions implied substantial downgrades both to GDP (Panel a) and in particular to the growth of consumption (Panel b) in the first years of the recovery, from 2.5% to 1.6%, and instead allocated much of the growth in GDP during the recovery to inventory accumulation. The estimate of long-run growth produced by our model is downgraded sharply as information about these revision is coming in, reflecting the role of consumption as the most persistent and forward looking component of GDP. This is clearly visible in Panels (c) and (d) of Figure 4.23. In particular, Panel (c) presents the evolution of the GDP nowcast for 2010 produced by Model 2 (black line), in comparison with the counterfactual nowcast that would result if there had been no revisions to long-run growth (dashed line). It is evident that the bulk of the revisions to GDP growth that year are the consequence of a large downward revision to long-run growth. Panel (d) plots the annual nowcast of GDP produced by Model 0 (dashed line), which does not allow for changes in long-run GDP growth, and Model 2 (solid line), our baseline specification. Up to mid-2010, both models produce similar nowcasts (not shown). After 2010, however, it is clearly visible that Model 0 begins each year expecting robust growth of above 3%, only to be disappointed by incoming data. The nowcasts by Model 2, which has incorporated the decline in long-run growth, do not suffer from the same pattern of systematic downward surprises.

Figure 4.23: Case Study: Impact of National Accounts Revision

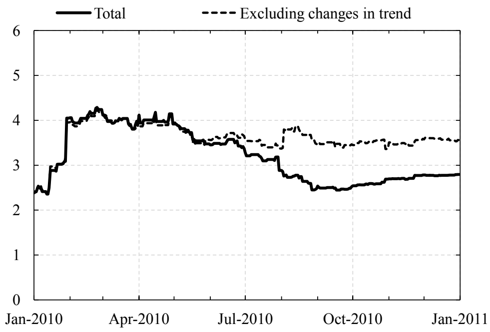
(a) Vintages of Real GDP



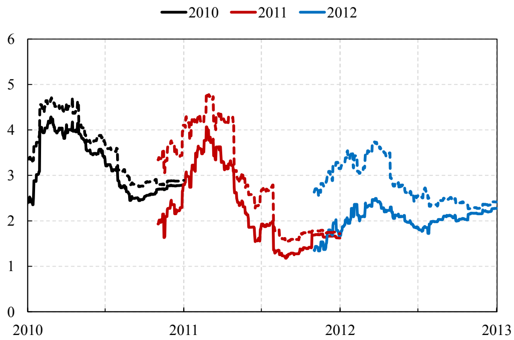
(b) Vintages of Real Pers. Consumption Exp.



(c) Breakdown of Model 2's nowcast for 2010



(d) Evolution of calendar-year GDP nowcasts



Note: Panels (a) and (b) compare several vintages of data on real GDP and real personal consumption expenditures around the time of important revisions by the BEA. Panel (c) presents the evolution of the GDP nowcast for 2010 produced by Model 2 (black line), in comparison with the counterfactual nowcast that would result if there had been no revisions to long-run growth (dashed line). The evolution of calendar-year nowcasts of real GDP growth produced by Model 0 (dashed) and Model 2 (solid) are presented in Panel (d).

4.7.11 Inspecting Data Set Size and Composition - More Details

Extended Model: Estimation Using a Very Large Panel

With regards to the size of the data set, in Section 4.4.1 of the main text we argue in favor of excluding disaggregated series within the various categories of real activity. This is because of the fact that the strong correlation across series within the same category might be a source of model misspecification. This is for two reasons: first, strong correlation in the idiosyncratic terms of series between the same category, and second, the fact that finer disaggregation levels are available for certain categories can lead to oversampling, see Boivin and Ng (2006) and Alvarez et al. (2012) for more details.

It is possible, however, to consider a more general specification of our model that can alleviate this problem, once we take into account the fact that persistent idiosyncratic movements common across series of the same category usually reflect differences in phase relative to the common activity factor. For example, all series related to employment respond to innovations to real activity with a lag. An interesting question is how our results are affected if one were to aim to make the dimension of $\mathbf{y}_t$ as large as possible, instead of carefully making variable selection based on the criteria discussed in the paper. In order to illustrate this point, we construct a "universe" of potentially available real activity time series for inclusion, based on a systematic attempt to find as many as possible US real activity time series. This is the "extended model" introduced in Section 4.4.6 of the paper.

Construction of the Extended Data Set

To construct the data panel for the extended model, we proceed as follows. First, we obtain all of the monthly real activity variables contained in the data set used by Stock and Watson (2012), which results in 75 time series.[54] Second, we exhaustively expand the monthly series contained in our original data set along all levels and dimensions of disaggregation available through Haver Analystics.[55] Out of this collection of expanded series, we then select any series that is not already contained in the 75 Stock and Watson indicators. Overall, this procedure results in a data set of as many as 155 time series capturing US real activity.[56]

Extended Model Specification

Maintaining the specification with a single factor (i.e. $k = 1$) we modify equation (4.1) of the paper as follows:

$$\mathbf{y}_t = \mathbf{c}_t + \mathbf{\Lambda}(L) f_t + \mathbf{u}_t, \tag{4.15}$$

[54]Details on this data set can be found in the online supplement to Stock and Watson (2012), available on Mark Watson's website. The only variable we were not able to obtain is Construction Contracts, which is not publicly available.

[55]This includes for example disaggregation along sectoral, regional and demographic characteristics.

[56]A detailed list of variables is available upon request.

such that the loading matrix $\Lambda(L)$ is now a polynomial in the lag operator of order s, i.e. contains the loadings on the contemporaneous common factor and its lags. In the special case where $s = 0$ we obtain our baseline specification. For the extended model, we set the maximum lag length, $s = 5$. The remaining equations of the model remain unchanged.

Priors and Model Settings

The data is standardized prior to estimation. "Minnesota"-style priors are applied to the coefficients in $\mathbf{\Lambda}(L)$, $\phi(L)$ and $\rho_i(L)$. More specifically:

- For the autoregressive coefficients of the factor dynamics, $\phi(L)$, the prior mean is set to 0.9 for the first lag, and to zero in subsequent lags. This reflects a belief that the common factor captures a highly persistent but stationary business cycle process.

- For the factor loadings, $\mathbf{\Lambda}(L)$, the prior mean is set to 1 for the first lag, and to zero in subsequent lags. This shrinks the model towards the factor being the cross sectional average of the variables, see D'Agostino et al. (2015).

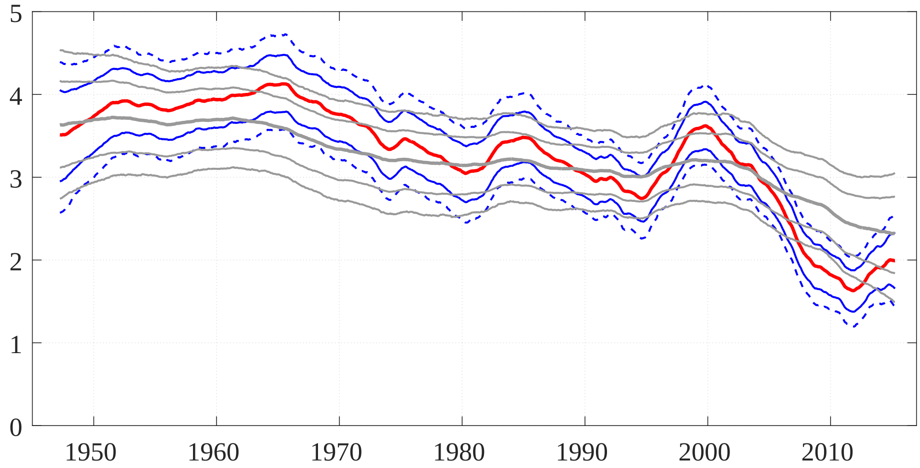
- For the autoregressive coefficients of the idiosyncratic, $\rho_i(L)$ the prior is set to zero for all lags, thus shrinking the model towards a model with no serial correlation in $u_{i,t}$.

In all cases, the variance on the priors is set to $\frac{\tau}{h^2}$, where τ is a parameter governing the tightness of the prior, and h is equal to the lag number of each coefficent, ranging $1:p$, $1:q$ and $1:s+1$. Following D'Agostino et al. (2015), we set $\tau = 0.2$, a value which is standard in the Bayesian VAR literature.

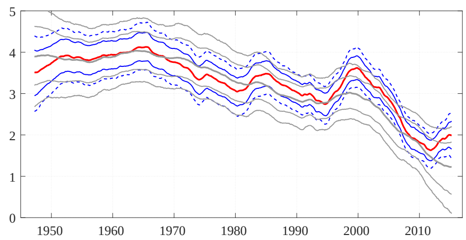
Results Across Alternative Data Sets

Figure 4.24: Comparison Across Alternative Data Sets/Models

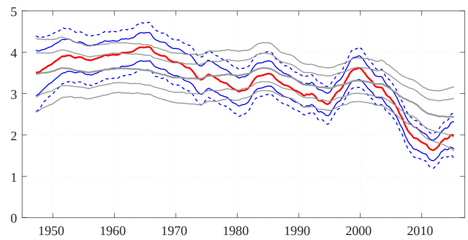
(a) Baseline With And Without Including Consumption



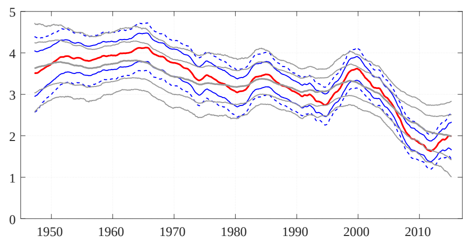
(b) Okun



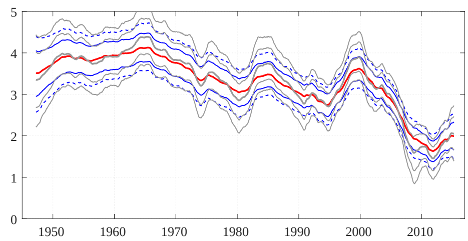
(c) Mariano-Murasawa



(d) Extended model ex. consumption



(e) Extended model



Note: In each panel our baseline the median estimate of real GDP growth is presented (red), with corresponding 68% (solid blue) and the 90% (dashed blue) posterior credible intervals. The corresponding estimates for the respective alternative data sets are superimposed in gray.

4.7.12 A Growth Accounting Exercise

The decomposition exercise carried out in Section 4.5 of the paper provides a first step towards giving an economically more meaningful interpretation of the movements in long-run real GDP growth detected by our model. While our equation $g_t = z_t + h_t$ follows from a simple identity, we demonstrate in this appendix how it can be related to the standard growth accounting framework.

To illustrate this point, consider two versions of the standard neoclassical growth model. In the first version, assume a standard Cobb-Douglas production function with Hicks-neutral technological change and constant returns to scale. In growth rates, this can be written as

$$dlogY_t = dlogTFP_t + \alpha dlogK_t + (1 - \alpha)dlogH_t, \tag{4.16}$$

where Y_t, K_t and H_t denote the level of output, the capital stock and labor input (total hours), respectively. α is the capital share and TFP_t is total factor productivity. Rearranging this relation gives

$$dlogY_t = dlogTFP_t + dlogH_t + \alpha(dlogK_t - dlogH_t), \tag{4.17}$$

so that the growth rate of real GDP is the sum of long-run growth in technology, total hours and a third term which captures differential growth in input factors which implies changes in the capital-labor ratio ("capital deepening"). In the second version of the neoclassical growth model, consider adding growth in labor-augmenting technology in the form of labor quality, denoted Q_t. In this case, the relation between growth rates is rearranged to

$$dlogY_t = dlogTFP_t + dlogH_t + \alpha(dlogK_t - dlogH_t) + (1 - \alpha)dlogQ_t. \tag{4.18}$$

Both relations (4.17) and (4.18) can be captured in our econometric framework. We define the first four elements of our vector of observables $\mathbf{y}_t$ in equation (4.1) to be the growth rate in real GDP, real consumption, TFP and total hours worked. As in the baseline model, *transitory* fluctuations in inputs (due to temporary shocks) would still be captured by the cyclical dactor, f_t, whereas the various sources of *permanent* changes in the growth rate of inputs (say, the long-run growth rate of technology, or the long-run growth rate of the population) would be included in $\mathbf{a}_t$. In order to mimic the relations prescribed by the two versions of the neoclassical growth model, we specify the long-run time variation in our model, $\mathbf{a}_t$ as a composite of three terms. While $\tilde{h}_t$ captures long-run movements in hours, the movements in long-run labor productivity are now further decomposed into a "technology" term $\tilde{z}_t$ and a "non-technology" term $\tilde{x}_t$. Formally, $\mathbf{c}_t$ in equation (4.2) is constructed as

$$\mathbf{a}_t = \begin{bmatrix} \tilde{z}_t \\ \tilde{h}_t \\ \tilde{x}_t \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{4.19}$$

What the non-technology term corresponds to depends on the underlying structure that is assumed. For instance, in the first version of the neoclassical growth model above

$$\tilde{x}_t \equiv \alpha(dlogK_t - dlogH_t) \tag{4.20}$$

and in the second case

$$\tilde{x}_t \equiv \alpha(dlogK_t - dlogH_t) + (1-\alpha)dlogQ_t. \tag{4.21}$$

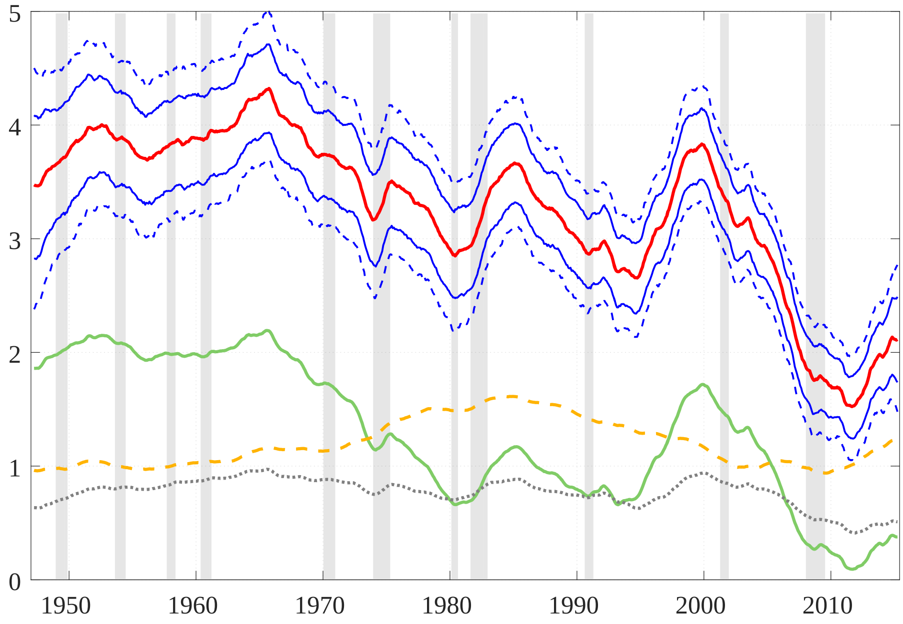
In both cases, $\tilde{x}_t$ subsumes potential long-run factors other than TFP that may explain changes in the long-run labor productivity trend we discuss in the paper.[57] $\tilde{z}_t$ is intended to capture changes in the long-run technology growth rate.

Figure 4.25 presents the results for US data when defining $\mathbf{c}_t$ by (4.19), and the measure of utilization-adjusted TFP from Fernald (2012) is used as an additional observable. Panel (a) shows the posterior estimate of long-run real GDP growth (including bands), together with the decomposition into long-run total hours growth, long-run technology growth and long-run non-technological growth. Reassuringly, the evolution of the total long-run growth component, g_t (red) is virtually identical to the estimate from our baseline model. The estimate of long-run hours growth (orange) is also very similar to its counterpart in Section 5 of the paper. Interestingly, the non-technological term (dashed gray) is positive on average and is relatively stable over the sample. Finally, the key insight from panel (a) is that our estimate of the long-run technology (green) displays strong movements that are very similar to the long-run growth rate of labor productivity which we have extracted in the simpler decomposition in Section 4.5. Under the assumption of a neoclassical structure, changes in long-run technology growth appear to be the main driver behind the recent slowdown in long-run real GDP growth. Panel (b) plots the growth rate of the utilization-adjusted TFP measure by Fernald (2012) in black, together with its long-run counterpart as estimated by our model (green, with blue bands). It is visible that the DFM approach is capable of extracting a smooth low frequency trend from the volatile series of TFP, which captures well-known episodes such as the 1970's slowdown and the IT boom of the 1990's. Overall, our framework is capable of providing an interesting angle on real-time movements in technology trends.

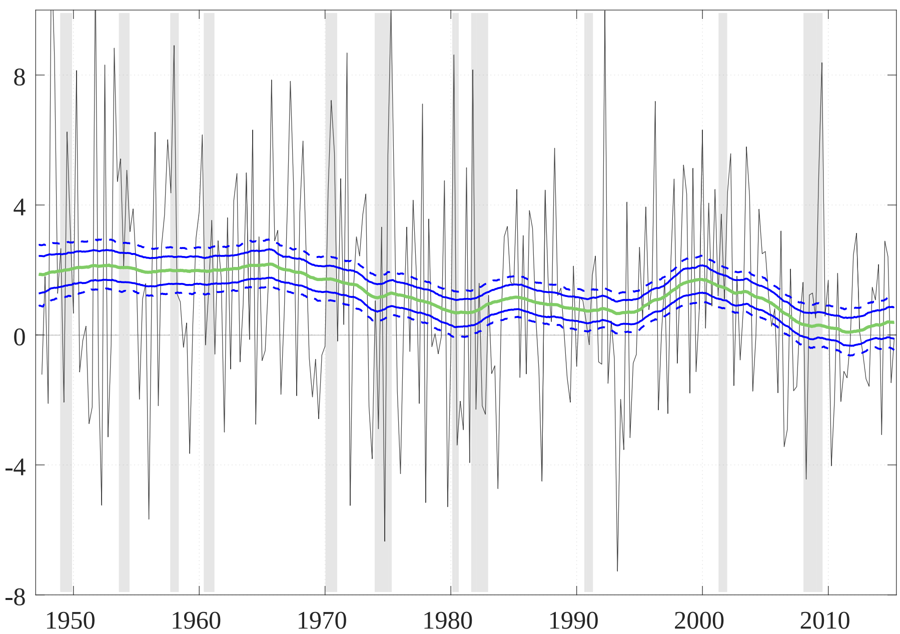
[57] Note here that in the first case we could also directly capture the technological parameter α into the matrix $\mathbf{B}$ by setting its (1,4) and (2,4) elements to α and interpreting $\tilde{z}_t$ directly as capital deepening. The specification above is somewhat more appealing in that it allows for a non-constant capital share. One can easily impose a constant value for α by scaling the posterior estimate of $\tilde{x}_t$ by that value.

Figure 4.25: Results of Growth Accounting Exercise

(a) Further Decomposition of US Long-Run US Output Growth



(b) Fernald (2012) TFP Growth: Data and Long-Run Estimate



Note: Panel (a) displays the posterior median estimates of long-run real GDP growth in red, together with the posterior median estimates of its components, long-run hours growth, long-run TFP growth and long-run non-technological growth (orange dashed, green, gray dotted). For long-run real GDP growth the corresponding with corresponding 68% and 90% posterior credible intervals are shown in solid and dashed blue. Panel (b) plots the growth rate of utilization-adjusted TFP by Fernald (2012) in black, together with its long-run counterpart in our econometric framework, i.e. the estimate of $\tilde{z}_t$, with corresponding 68% and 90% posterior credible intervals (green/blue).

4.8 Bibliography

Alvarez, R., M. Camacho, and G. Perez-Quiros (2012): "Finite sample performance of small versus large scale dynamic factor models," CEPR Discussion Papers 8867, C.E.P.R. Discussion Papers.

Auerbach, A. J. (2011): "Long-term fiscal sustainability in major economies," BIS Working Papers 361, Bank for International Settlements.

Bai, J. and P. Perron (1998): "Estimating and testing linear models with multiple structural changes," *Econometrica*, 66, 47–68.

Bai, J. and P. Wang (2015): "Identification and Bayesian Estimation of Dynamic Factor Models," *Journal of Business & Economic Statistics*, 33, 221–240.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2012): "Now-Casting and the Real-Time Data Flow," Working Papers ECARES 2012-026, ULB.

Banbura, M., D. Giannone, and L. Reichlin (2010): "Nowcasting," CEPR Discussion Papers 7883, C.E.P.R. Discussion Papers.

Banbura, M. and M. Modugno (2014): "Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data," *Journal of Applied Econometrics*, 29, 133–160.

Basistha, A. and R. Startz (2008): "Measuring the NAIRU with Reduced Uncertainty: A Multiple-Indicator Common-Cycle Approach," *The Review of Economics and Statistics*, 90, 805–811.

Benati, L. (2007): "Drift and breaks in labor productivity," *Journal of Economic Dynamics and Control*, 31, 2847–2877.

Beveridge, S. and C. R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle," *Journal of Monetary Economics*, 7, 151–174.

Bloom, N. (2014): "Fluctuations in Uncertainty," *Journal of Economic Perspectives*, 28, 153–76.

Boivin, J. and S. Ng (2006): "Are more data always better for factor analysis?" *Journal of Econometrics*, 132, 169–194.

Carter, C. K. and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

CBO (2001): "CBO's Method for Estimating Potential Output: An Update," Working papers, The Congress of the United States, Congressional Budget Office.

Clark, P. K. (1987): "The Cyclical Component of U.S. Economic Activity," *The Quarterly Journal of Economics*, 102, 797–814.

Clark, T. E. and M. W. McCracken (2012): "In-sample tests of predictive ability: A new approach," *Journal of Econometrics*, 170, 1–14.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices," *The Quarterly Journal of Economics*, 109, 241–65.

Cogley, T. (2005): "How fast can the new economy grow? A Bayesian analysis of the evolution of trend growth," *Journal of Macroeconomics*, 27, 179–207.

Cogley, T. and T. J. Sargent (2005): "Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S," *Review of Economic Dynamics*, 8, 262–302.

Creal, D., S. J. Koopman, and E. Zivot (2010): "Extracting a robust US business cycle using a time-varying multivariate model-based bandpass filter," *Journal of Applied Econometrics*, 25, 695–719.

D'Agostino, A., D. Giannone, M. Lenza, and M. Modugno (2015): "Nowcasting Business Cycles: A Bayesian Approach to Dynamic Heterogeneous Factor Models," Tech. rep.

de Wind, J. and L. Gambetti (2014): "Reduced-rank time-varying vector autoregressions," CPB Discussion Paper 270, CPB Netherlands Bureau for Economic Policy Analysis.

Del Negro, M. and C. Otrok (2008): "Dynamic factor models with time-varying parameters:measuring changes in international business cycles," NYFed Staff Rep. 326.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998): "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 863–83.

Diebold, F. X. and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13, 253–63.

Durbin, J. and S. J. Koopman (2012): *Time Series Analysis by State Space Methods: Second Edition*, Oxford University Press.

Edge, R. M., T. Laubach, and J. C. Williams (2007): "Learning and shifts in long-run productivity growth," *Journal of Monetary Economics*, 54, 2421–2438.

Evans, M. D. D. (2005): "Where Are We Now? Real-Time Estimates of the Macroeconomy," *International Journal of Central Banking*, 1.

Faust, J. and J. H. Wright (2013): "Forecasting Inflation," in *Handbook of Economic Forecasting*, ed. by G. Elliott and A. Timmermann, vol. 2.

Fernald, J. (2012): "A quarterly, utilization-adjusted series on total factor productivity," *Federal reserve bank of San Francisco working paper*, 19, 2012.

——— (2014): "Productivity and Potential Output Before, During, and After the Great Recession," *NBER Macroeconomics Annual 2014*, 29.

Fernald, J. G. and C. I. Jones (2014): "The Future of US Economic Growth," *American Economic Review*, 104, 44–49.

Fleischman, C. A. and J. M. Roberts (2011): "From many series, one cycle: improved estimates of the business cycle from a multivariate unobserved components model," Finance and Economics Discussion Series 2011-46, Fed Board of Governors.

Forni, M., D. Giannone, M. Lippi, and L. Reichlin (2009): "Opening The Black Box: Structural Factor Models With Large Cross Sections," *Econometric Theory*, 25, 1319–1347.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2003): "Do financial variables help forecasting inflation and real activity in the euro area?" *Journal of Monetary Economics*, 50, 1243–1255.

Gadea-Rivas, M. D., A. Gomez-Loscos, and G. Perez-Quiros (2014): "The Two Greatest. Great Recession vs. Great Moderation," WP Series 1423, Bank of Spain.

Geweke, J. (1977): "The Dynamic Factor Analysis of Economic Time Series," in *Latent Variables in Socio-Economic Models*, North-Holland.

Giacomini, R. and H. White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578.

Giannone, D., L. Reichlin, and L. Sala (2006): "VARs, common factors and the empirical validation of equilibrium business cycle models," *Journal of Econometrics*, 132, 257–279.

Giannone, D., L. Reichlin, and D. Small (2005): "Nowcasting GDP and inflation: the real-time informational content of macroeconomic data releases," Tech. rep.

——— (2008): "Nowcasting: The real-time informational content of macroeconomic data," *Journal of Monetary Economics*, 55, 665–676.

Gneiting, T. and A. E. Raftery (2007): "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378.

Gneiting, T. and R. Ranjan (2011): "Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules," *Journal of Business & Economic Statistics*, 29, 411–422.

Goldin, C. (2006): "The Quiet Revolution That Transformed Women's Employment, Education, and Family," *AEA Papers and Proceedings*, May 2006.

Gordon, R. J. (2004): "Why was Europe Left at the Station When America's Productivity Locomotive Departed?" Working Paper 10661, National Bureau of Economic Research.

——— (2010): "Okun's Law and Productivity Innovations," *American Economic Review*, 100, 11–15.

——— (2014a): "A New Method of Estimating Potential Real GDP Growth: Implications for the Labor Market and the Debt/GDP Ratio," NBER Working Paper 20423.

——— (2014b): "The Demise of U.S. Economic Growth: Restatement, Rebuttal, and Reflections," NBER Working Papers 19895.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87, 342–62.

Groen, J. J. J., R. Paap, and F. Ravazzolo (2013): "Real-Time Inflation Forecasting in a Changing World," *Journal of Business & Economic Statistics*, 31, 29–44.

Hansen, B. E. (1992): "Testing for parameter instability in linear models," *Journal of Policy Modeling*, 14, 517–533.

Harvey, A. C. (1985): "Trends and Cycles in Macroeconomic Time Series," *Journal of Business & Economic Statistics*, 3, 216–27.

Harvey, A. C. and J. H. Stock (1988): "Continuous time autoregressive models with common stochastic trends," *Journal of Economic Dynamics and Control*, 12, 365–384.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the equality of prediction mean squared errors," *International Journal of Forecasting*, 13, 281–291.

Hayashi, F. and E. C. Prescott (2002): "The 1990s in Japan: A Lost Decade," *Review of Economic Dynamics*, 5, 206 – 235.

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

Jacquier, E., N. G. Polson, and P. E. Rossi (2002): "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 20, 69–87.

Jorgenson, D. W., M. S. Ho, and K. J. Stiroh (2006): "Potential Growth of the U.S. Economy: Will the Productivity Resurgence Continue?" *Business Economics*, 41, 7–16.

Jurado, K., S. C. Ludvigson, and S. Ng (2014): "Measuring Uncertainty," *American Economic Review, Forthcoming.*

Kim, C.-J. and C. R. Nelson (1999a): "Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle," *The Review of Economics and Statistics*, 81, 608–616.

——— (1999b): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, The MIT Press.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–93.

Luciani, M. and L. Ricci (2014): "Nowcasting Norway," *International Journal of Central Banking, Forthcoming.*

Luo, S. and R. Startz (2014): "Is it one break or on going permanent shocks that explains U.S. real GDP?" *Journal of Monetary Economics*, 66, 155–163.

Marcellino, M., M. Porqueddu, and F. Venditti (2014): "Short-term GDP forecasting with a mixed frequency dynamic factor model with stochastic volatility," *Journal of Business & Economic Statistics, Forthcoming.*

Mariano, R. S. and Y. Murasawa (2003): "A new coincident index of business cycles based on monthly and quarterly series," *Journal of Applied Econometrics*, 18, 427–443.

McConnell, M. M. and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed since the Early 1980's?" *American Economic Review*, 90, 1464–1476.

Morley, J. C., C. Nelson, and E. Zivot (2003): "Why Are the Beveridge-Nelson and Unobserved-Components Decompositions of GDP so Different?" *The Review of Economics and Statistics*, 85, 235–243.

Nelson, C. R. and C. I. Plosser (1982): "Trends and random walks in macroeconomic time series: Some evidence and implications," *Journal of Monetary Economics*, 10, 139–162.

Nordhaus, W. (2004): "Retrospective on the 1970s Productivity Slowdown," Nber working papers, National Bureau of Economic Research, Inc.

Nyblom, J. (1989): "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223–230.

Ohanian, L. E. and A. Raffo (2012): "Aggregate hours worked in OECD countries: New measurement and implications for business cycles," *Journal of Monetary Economics*, 59, 40–56.

Oliner, S. D. and D. E. Sichel (2000): "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives*, 14, 3–22.

Orphanides, A. (2003): "The quest for prosperity without inflation," *Journal of Monetary Economics*, 50, 633–663.

Pain, N., C. Lewis, T.-T. Dang, Y. Jin, and P. Richardson (2014): "OECD Forecasts During and After the Financial Crisis: A Post Mortem," OECD Economics Department Working Papers 1107, OECD Publishing.

Perron, P. and T. Wada (2009): "Let's take a break: trends and cycles in US real GDP," *Journal of Monetary Economics*, 56, 749–765.

Pessoa, J. P. and J. Van Reenen (2014): "The UK Productivity and Jobs Puzzle: Does the Answer Lie in Wage Flexibility?" *The Economic Journal*, 124, 433–452.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

Pritchett, L. and L. H. Summers (2014): "Asiaphoria Meets Regression to the Mean," NBER Working Papers 20573, National Bureau of Economic Research, Inc.

Reifschneider, D., W. Wascher, and D. Wilcox (2013): "Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy," Finance and Economics Discussion Series 2013-77, Board of Governors of the Fed.

Reinhart, C. M. and K. S. Rogoff (2009): "The Aftermath of Financial Crises," *American Economic Review*, 99, 466–72.

Sargent, T. J. and C. A. Sims (1977): "Business cycle modeling without pretending to have too much a priori economic theory," Working Papers 55, Federal Reserve Bank of Minneapolis.

Sims, C. A. (2012): "Comment to Stock and Watson (2012)," *Brookings Papers on Economic Activity*, Spring, 81–156.

Stock, J. and M. Watson (2012): "Disentangling the Channels of the 2007-09 Recession," *Brookings Papers on Economic Activity*, Spring, 81–156.

Stock, J. H. and M. W. Watson (1989): "New Indexes of Coincident and Leading Economic Indicators," in *NBER Macroeconomics Annual 1989, Volume 4*, National Bureau of Economic Research, Inc, NBER Chapters, 351–409.

——— (1999): "Business cycle fluctuations in us macroeconomic time series," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1 of *Handbook of Macroeconomics*, chap. 1, 3–64.

——— (2002): "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

——— (2003): "Has the Business Cycle Changed and Why?" *NBER Macroeconomics Annual 2002*, 17, 159–230.

Summers, L. (2014): "Secular Stagnation," IMF Economic Forum: Policy Responses to Crises, speech delivered at the IMF Annual Research Conference, November 8th.

Whelan, K. (2003): " A Two-Sector Approach to Modeling U.S. NIPA Data," *Journal of Money, Credit and Banking*, 35, 627–56.