Essays on Information Frictions and the Macroeconomy

by

Andras Komaromi

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor David H. Romer, Chair
Professor Yuriy Gorodnichenko
Professor Demian G. Pouzo
Professor James A. Wilcox

Spring 2015

Essays on Information Frictions and the Macroeconomy

Abstract

Essays on Information Frictions and the Macroeconomy

by

Andras Komaromi

Doctor of Philosophy in Economics

University of California, Berkeley

Professor David H. Romer, Chair

This dissertation is a compilation of three essays on the role of information frictions in macroeconomics.

The *first essay* contributes to the literature on the impact of uncertainty on the business cycle. The cross-sectional dispersion of firm-level outcomes, such as sales growth or stock returns, is markedly countercyclical. Recent papers have framed this fact as evidence that exogenous "uncertainty shocks" are important drivers of business cycles. This paper provides empirical evidence that the co-movement of various dispersion measures with the business cycle is better understood as the economy's endogenous response to traditional first moment shocks – dispersion is the effect, not the cause. It then develops a theoretical model that links the cross-sectional dispersion of micro-level outcomes to the aggregate state of the economy. The mechanism is based on time-varying rational inattention. In bad times, firms pay more attention to idiosyncratic shocks hitting their business environment. More precise micro-level information about the underlying heterogeneity leads to higher dispersion in realized outcomes. In line with the empirical findings, the model generates countercyclical dispersion without relying on exogenous second moment (uncertainty) shocks.

The *second essay* uses survey expectations to assess the microfoundations of an important class of macroeconomic models. Many theoretical macro models try to explain the pervasive nominal and real stickiness in the data by assuming rational decision-making under imperfect information. The behavior of consensus (average) forecasts is consistent with the predictions of these models, which can be seen as supportive empirical evidence for the models' microfoundations (Coibion and Gorodnichenko, 2012). This paper demonstrates, however, that the individual-level data underlying the consensus forecasts are at odds with this interpretation. In particular, I document that individual expectations in the Survey of Professional Forecasters do not pass a very weak test of rational expectations: current forecast revisions are strong predictors of subsequent forecast errors. Information frictions alone cannot explain this pattern.

I go on to propose a simple modification of the noisy information framework that allows for a particular form of non-rational expectations: agents may incorrectly weight new infor-

mation against their prior. I show that this parsimonious model can match the survey data along several dimensions. Using the structure of the model, I estimate the direction and size of inefficiencies in the expectations formation process. I find that in most cases agents put too much weight on their private information, which can be interpreted as overconfidence in the precision of private information. I also show that there is substantial heterogeneity across agents in the deviation from rational expectations, and I relate these differences to observable characteristics. Finally, I discuss potential interpretations of my empirical results and their implications for macroeconomic theory.

The *third essay* explores the potential trade-off between competition and systemic stability in financial intermediation. Why do banks feel compelled to operate with such high leverage despite the risks this poses? Using a simple model, I argue that the degree of competition goes a long way in explaining capital structure decisions. On the one hand, information frictions (adverse selection) render debt a cheaper form of financing than equity. On the other hand, more reliance on debt increases the probability of bankruptcy, which results in the loss of the bank's charter value. The degree of competition affects charter values, and hence changes the way banks balance between these two forces. A panel analysis of European banks' capital structure around the introduction of the euro reveals statistically and economically significant effects consistent with this hypothesis. Banks, in particular smaller banks, decreased their equity ratios after entering the currency area. Complementary evidence suggests that this effect can be attributed to increased competitive pressures boosted by the euro.

To Szilvia and the Beautiful Baby Girl in Her Womb

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my deepest gratitude to Professor David Romer, Professor Yuriy Gorodnichenko and Professor Christina Romer for their continuous support and guidance. They provided me indispensable encouragement when I stumbled. I also thank Professor Pierre-Olivier Gourinchas for helpful discussions and for his faith in me. John Mondragon and Hedvig Horvath were wonderful friends and colleagues over the years.

Preface

The core of neoclassical economic theory was built around the idea of fully informed rational consumers and firms interacting in a well-functioning marketplace. The standard assumption about the information environment encompasses three important elements. First, agents have perfect information, which means that they have perfect knowledge of all relevant prices, utilities, production methods, policy decisions and other parameters of the economy. Second, agents have rational expectations, which means that they are able to combine efficiently all their information to form mathematically correct expectations about the future. Third, information is symmetric, that is no one party participating in a transaction has more or better information than the other. These assumptions are necessary building blocks of the powerful theoretical structure in which neoclassical economists demonstrated some groundbreaking results, such as the fundamental theorems of welfare.

For a long time, however, economists have recognized that these assumptions are not generally applicable when we try to understand certain market situations or fluctuations in the aggregate economy. Information is often not perfect; either because it is impossible to observe all relevant economic variables and relationships or because doing so would be too costly compared to the potential gains from better information. Formal models of imperfect information explicitly acknowledge that consumers and firms have to make decisions based on outdated or noisy information. Similarly, economic agents may fail the high bar of rational expectations when making predictions about uncertain outcomes. Behavioral models of learning and expectations formation have emphasized the role of simple heuristics and rule of thumbs when people process information. Finally, a large body of research has studied the implications of asymmetric information. The theoretical concepts of adverse selection, moral hazard, screening and signaling have been used successfully in numerous applied contexts such as in labor economics, health economics and corporate finance.

Fitting into this broad literature, one unifying theme of the essays in this dissertation is the emphasis on information frictions in explaining economic phenomena. I use the term *information friction* to refer to any deviation from the information environment described in the first paragraph. Each of the three essays documents some puzzling or otherwise interesting empirical facts, and then uses a theoretical framework with an information friction to explain the observations. The model of the first paper dispenses with the assumption of full information. Instead, it assumes that firms can choose endogenously how much information to acquire about their business environment, subject to the cost of information gathering.

I show that incorporating this information friction sheds new light on the relationship between the business cycle and common measures of uncertainty. The second paper relaxes the assumption of rational information processing. It provides evidence that survey data on peoples expectations are inconsistent with rational expectations macroeconomic theories, and proposes a simple behavioral model to interpret the deviations. The third paper highlights an important interaction between asymmetric information and market structure in the context of financial intermediation. Financial intermediaries prefer debt to equity financing due to adverse selection, but taking on more debt also increases the probability of default. Higher competitive pressure reduces the cost of default by lowering the charter value of intermediaries, and thus it may induce high leverage of individual banks that creates systemic instability in the financial sector.

The three essays in the dissertation also share a common methodological approach. Although my work on these topics involves some theorizing, each paper has a strong empirical and applied focus. I motivate my research questions with empirical facts, and I use disaggregated data and micro-to-macro analysis to shed light on economic problems. In Chapter 1, I use firm-level accounting and stock market data to study the relationship between the business cycle and the cross-sectional dispersion of micro-level outcomes. Chapter 2 investigates a rich dataset of individual survey forecasts to test the microfoundations of a large class of macroeconomic models. In Chapter 3, I analyze bank balance sheets around the introduction of the euro to understand whether individually optimal decisions in the face of increased competition can contribute to systemic instability at the aggregate level.

I am hopeful that my dissertation demonstrates that taking information frictions seriously and utilizing microdata is a fruitful approach to answer interesting questions in seemingly distant areas of macroeconomics.

Chapter 1

Attention and Dispersion over the Business Cycle

1.1 Introduction

The rational inattention literature has highlighted that economic actors' information processing capacity or attention is finite, so they have to choose how much and what kind of information to acquire before making decisions (Sims, 2003). An important insight is that even if all relevant information is available in principle, decision-makers with limited capacity will still behave as in imperfect information environments. However, the precision of agents' information becomes endogenous in this approach, which opens up a new channel for the transmission of exogenous shocks. This paper argues that the optimal attention allocated to idiosyncratic conditions varies over the business cycle, which affects the cross-sectional moments of economic variables. First and foremost, endogenous information choice provides an alternative explanation of the stylized facts that have gained a central role in the recent literature on "uncertainty" shocks.

The cross-sectional dispersion of economic variables such as labor income, prices, sales and stock returns is markedly countercyclical. How should we interpret this fact? In the past few years, as uncertainty has become a central topic in policy discussions and in academic research, a number of papers have relied on cross-sectional dispersion measures to proxy for micro-level uncertainty. Most prominently, Bloom (2009) argues that rising dispersion of firm performance in recessions is *prima facie* evidence for higher firm-level uncertainty in bad times. According to this view, causality runs from dispersion to aggregate economic activity. Changes in cross-sectional dispersion reflect exogenous changes in the volatility of micro-level economic shocks, which in turn impact aggregate output through non-convex adjustment costs, irreversibilities or financial market frictions.

In this paper I challenge this interpretation of countercyclical dispersion, and I propose an alternative mechanism inspired by the rational inattention literature. In my theory, the second moment of idiosyncratic shocks is constant, but the attention allocated to these shocks changes over the business cycle. Since information choice governs the covariance between choice variables (for example, prices) and the relevant states (for example, firm-level business conditions), time-varying attention has the potential to generate changes in dispersion over time. Endogenous information choice has been used extensively in macroeconomics to explain agents' incomplete reaction to exogenous shocks (Reis, 2006; Mackowiak and Wiederholt, 2009), but these papers adopt a timeless perspective for the attention allocation problem. Here I study how the incentive to acquire information may change according to the aggregate state of the economy.

My empirical work demonstrates that standard dispersion measures do not merely reflect exogenous second moment shocks. A simple argument can be based on the timing and frequency of spikes in cross-sectional dispersion. In particular, measures of micro-level dispersion tend to lag the cycle, and they rise sharply in nearly all recessions for which data are available. Since both conventional wisdom and VAR-based historical decompositions of the business cycle indicate that it is not the case that uncertainty shocks played a major role in every recession, this is suggestive of a causal effect from recessions to micro-level dispersion.

Formal statistical analysis confirms that dispersion reacts endogenously to output fluctuations, and moreover, that this response is the sole source of the negative co-movement between output and the dispersion series. Disentangling the role of second moment and traditional first moment shocks in the joint evolution of output and dispersion is not straightforward, specifically because we lack an accepted measure of exogenous changes in idiosyncratic uncertainty. Papers studying the macroeconomic role of uncertainty have relied on a triangular orthogonalization of VAR innovations, which depend on arbitrary assumptions and do not use any external information that we might have from other sources about the drivers of the business cycle. I modify previous identification schemes in two steps. First, I collect an array of first moment shocks that have been identified in the business cycle literature and are thought to capture exogenous variation in aggregate economic activity, such as oil price shocks and monetary policy surprises. Second, I use these shock series as "external instruments" in a structural VAR to uncover the response of dispersion to an output level-shock (Stock and Watson, 2008; Mertens and Ravn, 2013). This overidentified GMM methodology allows me to incorporate all the information that is jointly contained in these shock series.

If idiosyncratic uncertainty – proxied by cross-sectional dispersion – is really an independent structural shock on its own right, then it should not respond to other aggregate shocks. However, if micro dispersion endogenously increases in bad times, we would expect that dispersion measures would go up following a negative aggregate shock. This is exactly what I find, which shows that usual indicators of micro-level uncertainty have a substantial endogenous component. Furthermore, identified dispersion innovations in my SVAR do not trigger the negative output effects that uncertainty-driven business cycle models imply. In fact, innovations in dispersion do not contribute at all to the negative co-movement of output and cross-sectional dispersion. These findings are in stark contrast with the findings of earlier studies that used simple recursive VARs to identify the effect of uncertainty shocks on output.

I also show that mechanical effects involving heterogeneous cyclical sensitivities among firms cannot explain the causal link from aggregate output to cross-sectional dispersion. First, removing the common component from stock returns and sales growth leaves the co-movement of output and dispersion virtually unchanged. Hence, different loadings on aggregate factors cannot account for my findings. Second, explanations based on some factor structure only work if we impose a certain form of asymmetry on the loadings. I show that these restrictions do not bear out in the data. I conclude that fluctuations in aggregate economic activity affect the dispersion of firm-level outcomes, and that this phenomenon cannot simply be attributed to the differential impact of aggregate shocks across firms.

Motivated by my empirical results, I propose a new mechanism to explain the co-movement of cross-sectional dispersion and aggregate output. The general idea is that recessions are times when incentives to acquire information about the inherent heterogeneities in the economy are high. Consequently, the observed jump in cross-sectional dispersion does not reflect an exogenous rise in underlying heterogeneity (a second moment shock), but rather is the result of increased awareness about differences that were hidden during the boom. When business conditions are generally good, the benefit from paying attention to micro-level differences is low, so little information production takes place and individual decisions converge on the average. A deterioration of the aggregate state of the economy (a first moment shock) increases the expected gain from better information, so agents put more effort into uncovering micro-level differences. As a result, these differences factor more strongly into agents' decisions.

My equilibrium model formalizes this general idea in the context of price-setting among heterogeneous firms. Firm heterogeneity is captured by different productivity levels, but this could represent other variation in the micro-level business environment, such as differing demand conditions. Crucially, firms cannot freely observe their idiosyncratic productivity. Instead, they have to sacrifice resources to acquire information in the form of a noisy signal, which then allows them to make better decisions. Business cycle fluctuations are driven by conventional first-moment TFP shocks that shift the mean of the (log) productivity distribution in the economy. I take a simple approach to generate countercyclical incentives to information acquisition. I model information gathering as a time-consuming process that requires labor input that could also be used in other productive activities. As a result, the cost of information acquisition tracks the real wage, which is low in bad times. However, I also discuss other mechanisms that can induce more learning in recessions. For example, more precise micro-level information may help firms avoid the costs of external financing. Since a negative aggregate shock pushes all firms closer to their external financing threshold, the expected benefit from better information increases in bad times.

Regardless of the source countercyclical information acquisition, if firms pay more attention to their idiosyncratic shocks when aggregate conditions are worse, then we will observe higher dispersion in firms' optimal choices in recessions. I show that my model provides a unifying interpretation for the countercyclical dispersion of four firm-level variables: (i) price changes, (ii) sales growth, (iii) real value-added growth, and (iv) stock returns. Furthermore, the mechanism is quite general, so it can be applied in different settings to shed light on the cross-sectional dispersion of other variables. For example, the countercyclical dispersion of labor income can be explained if employers make more effort to uncover idiosyncratic worker productivity in recessions.

An important difference between my mechanism and the models featuring shocks to idiosyncratic uncertainty is that they have completely opposing predictions about the amount of micro-level information over the business cycle. In an "exogenous uncertainty world" booms are periods of more precise information, and recessions are characterized by less precise knowledge. According to the "endogenous information view", information about micro heterogeneity decays in expansions, because decision-makers choose not to pay much attention to it. During downturns, the benefit of better information increases, so information gets replenished endogenously. From a policy perspective, understanding the source of time-varying dispersion seems important, as policies designed to reduce uncertainty almost certainly differ from policies designed to alter incentives for information acquisition.

Related literature. This paper builds on the now voluminous literature on the role of uncertainty in business cycles. I do not attempt to summarize this literature here, but refer the reader to a recent survey by Bloom (2014). The vast majority of these papers take an "exogenous uncertainty" view, that is they explore the aggregate implications of second moment shocks.[1] However, a smaller number of papers argue that higher measured micro-level uncertainty, often labeled "micro-volatility", can arise endogenously in recessions. For example, in D'Erasmo and Moscoso-Boedo (2011), firms can enter new markets after paying a fixed cost. Market participation is endogenous, and varies over the cycle. In booms, firms expand and end up diversifying market-specific demand shocks by servicing more markets. The mechanism in Alessandria et al. (2014) is also based on market participation, but in an open economy context. In their model, a negative country-specific shock affects non-exporters differently from exporters, leading to a reallocation of production across these heterogeneous producers and to rising dispersion. Ilut, Kehrig and Schneider (2014) study the distribution of employment growth when firms adjust asymmetrically to dispersed but correlated signals. If hiring decisions respond more to bad news than to good news, the cross sectional dispersion of employment growth is countercyclical. The authors show that this asymmetry naturally arises if firm decision-makers are ambiguity averse.

The mechanism in my paper is perhaps most closely related to Bachmann and Moscarini (2012). Firms in their model are unsure about their demand elasticity and they infer it from the response of their sales to price changes. Subject to a fixed operating cost, recessions are the best times to price-experiment, because the resulting information helps decide whether to exit the market. Although their explanation also emphasizes endogenous learning, it is different from this paper. In my story, active information gathering about time-varying idiosyncratic fundamentals leads to higher dispersion by moving the economy closer to the perfect information outcome. In Bachmann and Moscarini (2012), firms change prices and create dispersion *in order to* learn about a constant parameter. In addition to the difference in model mechanisms, my work also adds to the previous literature by empirically testing the direction of causality between dispersion and the business cycle.

[1]See Section 1.2 for some relevant papers from this literature.

Most of the work on imperfect information in macroeconomics is concerned with learning about aggregate shocks, in particular the stance of monetary policy. In contrast, my paper is interested in the attention allocated to idiosyncratic conditions, and how it affects cross-firm dispersion. David, Hopenhayn and Venkateswaran (2014) have a similar motivation: they introduce imperfect information about idiosyncratic shocks in order to explain resource misallocation among firms. My paper differs in two important ways. First, the precision of the private and public signals in their model is exogenous to the firm. My model has a simpler information structure, but the amount of information is a choice variable for the firm. Second, they do not have aggregate uncertainty. My goal is precisely to study the impact of aggregate fluctuations on learning about idiosyncratic shocks.

Finally, the model of attention spurred by a reduction in the opportunity cost of the inputs used in information acquisition is similar in spirit to business cycle theories about the "virtues of bad times" described by Aghion and Saint-Paul (1998) and Hall (1991). In these papers, firms take advantage of a negative shock by shifting resources into alternative uses, such as investing in innovation or reorganization. Recently, Bloom et al. (2014) revived this idea to demonstrate the possibility of large dynamic gains from trade opening. They argue that firms suddenly exposed to import competition from China had an incentive to shift their resources into innovation, because the opportunity cost of these inputs declined. My mechanism is reminiscent of this cleansing effect of recessions, but in this case firms' alternative use of time is information-gathering.

The organization of the paper is the following. Section 1.2 presents the empirical evidence that micro-volatility responds endogenously to output fluctuations. Section 1.3 shows that this result cannot be explained by mechanical effects arising from heterogeneous sensitivities to common shocks. Section 1.4 presents the model, whose implications for the cyclical properties of dispersion are analyzed in Section 1.5. Section 1.6 concludes. The Appendices contain robustness checks (A.1), the details of the econometric approach (A.2), an alternative motivation of countercyclical attention (A.3), and the model derivations (A.4).

1.2 Dispersion over the Business Cycle: Cause or Effect?

This section presents a series of tests for the direction of causality between aggregate fluctuations and cross-firm dispersion. The uncertainty literature views dispersion as a measure of idiosyncratic uncertainty, and argues that changes in uncertainty cause movements in aggregate output. I briefly summarize this view, and then I show that the interpretation of these time series as an independent structural shock process that drives business cycles is problematic. First, I present evidence that dispersion increases in all recessions, even in those unlikely to be driven by uncertainty shocks. Second, and more importantly, a SVAR analysis utilizing external instruments reveals that cross-sectional dispersion measures respond to changes in the level of aggregate output. These facts suggest that dispersion increases endogenously after some shock has pushed the economy into recession.

The uncertainty view of dispersion

To set the stage, it is useful to briefly discuss what the literature means by time-varying uncertainty and how it is related to changes in dispersion. Most papers adopt a framework similar to the one in Bloom et al. (2012). Here, the technology of firm i is described by a production function

$$y_{it} = A_t z_{it} f(k_{it}, n_{it}),$$

where each firm's productivity is a product of an aggregate component, A_t, and an idiosyncratic component, z_{it}. The technology terms follow

$$\begin{aligned} log(A_t) &= \rho log(A_{t-1}) + \sigma^A_{t-1}\varepsilon_t, \\ log(z_{i,t}) &= \rho log(z_{i,t-1}) + \sigma^Z_{t-1}\varepsilon_{it}. \end{aligned}$$

In this formulation, σ^A represents aggregate uncertainty and σ^Z governs micro-level uncertainty. The volatility of the idiosyncratic component, z_{it}, implies that productivity (or demand) dispersion *across firms* is time-varying, while volatility in the aggregate component, A_t, implies that *all firms* are affected by more volatile shocks. Consequently, these two shock processes should drive different statistics of the data. Increases in σ^A induce higher time series variability in aggregate variables like GDP growth or the S&P500 index, while changes in σ^Z imply that the *cross-sectional dispersion* of firm performance (for example, sales or stock market returns) is time-varying.

Recently, cross-sectional dispersion measures of micro-level uncertainty (micro-volatility) have received much attention in the literature. On the empirical front, Bloom et al. (2012) look at performance measures of US firms, and they find that the cross-sectional dispersion of these variables rises in recessions. Bachmann and Bayer (2013) document the same stylized fact for a panel of German firms. Gilchrist, Sim and Zakrajsek (2014) show that the dispersion of credit spreads on US corporate bonds is also countercyclical. Berger and Vavra (2010) find that the cross-sectional dispersion of item-level price changes in BLS microdata is countercyclical.

In line with the empirical literature, a number of theoretical papers completely abstract from changes in aggregate uncertainty, and model only the variation in σ^Z. These models treat the cross-sectional dispersion of economic variables as indicators of uncertainty, and rationalize them by introducing a stochastic process for micro-level uncertainty that can be hit by random shocks. With this interpretation, the research agenda is to find mechanisms through which micro uncertainty shocks (changes in σ^Z) can affect the level of output. The research question in Arellano, Bai and Kehoe (2012) is a very clear example of this interpretation: *"Can an increase in the volatility of firm-level idiosyncratic shocks that generates the observed increase in the cross-section dispersion in the recent recession lead to a sizeable contraction in aggregate economic activity?"* Similar interpretations are offered by, for example, Gilchrist, Sim and Zakrajsek (2014) and Bachmann and Bayer (2013).

Several mechanisms have been proposed to translate second moment shocks into first moment aggregate effects. These include risk aversion, financial contracting frictions, and non-convex adjustment costs and irreversibilities which can lead to a wait-and-see effect when firms are facing higher uncertainty. Not surprisingly, after taking a stand on the direction of causality, these models are able to match the co-movement of output and some of the dispersion measures.

But is it obvious that this is the stand we should take? I will argue that careful statistical analysis of the relationship between micro uncertainty indicators and the business cycle poses important challenges to the direction of causality assumed in the literature.

Instruments and data sources

This section describes the sources and the construction of my dataset. Both the macro and micro databases used in this paper are standard, so I keep the discussion to a minimum. I consider two measures advocated by Bloom (2009) and Bloom et al. (2012) as proxies for micro-level uncertainty: the cross-sectional dispersions of firm-level sales growth and firm-level stock returns. Sales figures come from the quarterly income statements of publicly traded US firms as reported in the Compustat database. I calculate annual growth rates to filter out seasonal patterns. Monthly stock returns of listed US firms are retrieved from the Center for Research in Security Prices (CRSP) database. I construct quarterly returns by summing up the monthly numbers. I use the interquartile range as a robust measure of cross-sectional dispersion. Following Bloom (2009), I only include firms with at least 25 years of data. This makes the sample much more homogeneous, and allows me to abstract from entry and exit, in line with the model presented later in the paper. In the graphs and statistical results reported here, I use the dispersion of the raw sales growth and stock return series. However, Appendix A.1 shows that removing the common component of firm-level variables leaves the dispersion time series virtually unchanged. I also demonstrate that using other cut-off values for inclusion in the sample (5, 10 and 15 years) produces very similar results.

The next section presents a historical decomposition of the business cycle based on a VAR with six macroeconomic variables (industrial production, employment, hours, CPI, hourly wage, Fed Funds rate) and two stock-market variables (S&P500 index, VIX index). These aggregate variables were downloaded from the FRED database.

The external instruments that I use to implement my identification scheme fall into five categories: monetary policy surprises, fiscal shocks, exogenous oil price increases, productivity shocks and credit supply shocks. I use these instruments in a structural VAR framework to isolate movements in output that are not caused by micro-level uncertainty shocks. In constructing my dataset, I draw on the large literature that tries to identify aggregate structural shocks. Stock and Watson (2012) provides a comprehensive list of papers that have constructed aggregate shocks to explain business cycles. I chose the twelve series with the longest sample period, making sure that each category is roughly equally represented. I also preferred shock series whose construction does not rely heavily on model-dependent assumptions, such as restrictions in a VAR, but rather on survey responses, financial market data or narrative evidence. The structural shocks are summarized here; specific sources, sample periods and some details are provided in Table 1.1.

Monetary policy shocks. I include three proxies for monetary policy shocks. The first is from Gürkaynak, Sack and Swanson (2005), who infer surprise changes in the target rate by looking at changes in the Federal Funds Future rate in narrow windows around FOMC announcements. The second proxy is from Coibion et al. (2012) who extend the Romer and Romer (2004) monetary policy shock series by incorporating new data that have become available. This series is computed as the residual from a regression of a constructed Federal Reserve monetary intentions measure on internal Federal Reserve forecasts. The third measure of monetary policy shocks comes from a VAR with short run restrictions as in Christiano, Eichenbaum and Evans (1999).

Fiscal policy shocks. I utilize three fiscal shock series: the Ramey (2011) federal spending news series, Fisher and Peters (2010) excess returns on stocks of military contractors, and the Romer and Romer (2010) exogenous tax change series.

Oil shocks. I use two oil shock series. The Hamilton (1996) monthly net oil price increase is the percentage amount by which the oil price in a month exceeds the previous peak over the past twelve months (constructed from the PPI for oil). I take quarterly averages of the monthly shocks. The Kilian (2008) oil shock is his measure of OPEC production shortfall stemming from wars and civil strife.

Productivity shocks. I use two productivity shock measures. The first is an updated version of the Basu, Fernald and Kimball (2006) series on quarterly total factor productivity adjusted for variations in factor utilization from Fernald's webpage. The second series is constructed using the Gali (1999) identification scheme based on long run restrictions. Specifically, technology shocks are identified from the restriction that only technology shocks have long run effect on productivity.

Credit supply shocks. I use two measures that capture exogenous changes in financial intermediation. The Gilchrist and Zakrajsek (2012) excess bond premium shocks try to isolate movements in corporate bond spreads that are orthogonal to the current state of the economy. The other measure is the bank loan supply shock from Bassett et al. (2012), which they compute as the unpredictable component of bank-level responses to the Fed's Senior Loan Officer Survey.

Suggestive evidence of endogenous dispersion

Figures 1.1 and 1.2 show the cross-sectional dispersion of quarterly sales growth and stock returns of publicly traded firms in the US. Shaded bars indicate NBER recession dates. It is immediately apparent from the graphs that micro-level dispersion rises sharply in all recessions. Figure 1.3 and 1.4 focus on the business cycle properties of dispersion. First, I plot the cyclical components of GDP and the dispersion measures using the Baxter and King (1999) bandpass filter. The negative co-movement is even more discernible at business cycle frequencies. I also report the time series correlation of the dispersion series and real

GDP growth, as well as the regression coefficient on a recession dummy. Both are negative and statistically significant.

What does this strong co-movement suggest about the causal link between aggregate economic activity and dispersion? If cross-sectional dispersion was solely driven by uncertainty shocks, one should conclude that uncertainty shocks featured prominently in every recession. This conclusion seems to be at odds with conventional wisdom and model-based decompositions of the business cycle. For example, in Figure 1.5 I use an 8-variable VAR from Bloom (2009) to calculate the historical contribution of various shocks to changes in industrial production. Bloom (2009) reports impulse response functions from this model to demonstrate the effects of time-varying uncertainty on economic activity. In his VAR, uncertainty is proxied by the VIX index, and structural shocks are obtained by Choleski factorization. According to my historical decomposition, the adverse effect of uncertainty is apparent in the Great Recession, but it is hard to see any systematic pattern in other recessions. To highlight this point, consider the 1981-82 recession. Most economists (and the VAR) agree that the fall in economic activity was caused largely by monetary policy. Yet, indicators of cross-sectional dispersion rose markedly.[2]

It is also evident from the figures that both sales growth dispersion and stock return dispersion tend to peak towards the end of recessions.[3] To formally test this hypothesis, Table 1.2 reports the results from Granger causality tests between the output gap and the dispersion indicators. The results confirm the visual impression that, on average, the business cycle leads changes in dispersion.

Based on the graphs and the lead-lag relationships in the data, it appears hard to make a case that exogenous second moment shocks, as reflected in the dispersion-based uncertainty indicators, lead to recessions. Rather, the timing of events is consistent with an interpretation where some aggregate shock pushes the economy into recession, which then leads to elevated levels of dispersion. This interpretation will be studied more carefully in the next section.

Dispersion and first moment shocks

To identify the impact of business cycle fluctuations on dispersion, I isolate exogenous movements in output induced by standard first-moment shocks. If uncertainty – measured by cross-sectional dispersion – is really an independent structural shock on its own right, then it should not respond to other aggregate shocks. However, if micro dispersion endogenously increases in bad times, we would expect our dispersion measures to go up following a negative aggregate shock. I implement my test in two steps. First, I estimate the impulse response function of dispersion measures to each first-moment shock separately. Second, after con-

[2] Without doubt, the triangular identification scheme in the Bloom (2009) VAR is ad hoc. The only goal of this exercise is to show that even the methodology of this seminal paper suggests that rising dispersion in recessions cannot be fully explained by uncertainty shocks.

[3] The one clear exception is the 2001 episode, when stock return dispersion reached its highest level right at the beginning of the recession. Presumably this is related to the fact that this downturn followed the collapse of the Dot-com bubble in 2000.

firming that individual shocks do affect the uncertainty indicators, I carry out a structural VAR analysis between aggregate output and cross-sectional dispersion. In this exercise, I use multiple shock series together to filter out the effect of exogenous output changes on dispersion.

Using the exogenous shocks directly

I estimate simple time series models to check whether dispersion measures can be predicted by each of the structural shocks. More precisely, I run the following regressions:

$$d_t = a + \sum_{k=1}^{p} b_k FMS_{t-k} + \sum_{j=1}^{q} c_j d_{t-j} + \varepsilon_t \tag{1.1}$$

where d_t is either sales growth dispersion or stock return dispersion, and FMS_t is one of the first-moment aggregate shock series. This specification forces the contemporaneous response of dispersion to aggregate shocks to be zero, so it can be regarded as a conservative benchmark. However, the results are practically unchanged if I include FMS_t on the right-hand side. In the reported results I set $p = q = 4$.

Figure 1.6, 1.7, 1.8, 1.9, and 1.10 plot the implied impulse responses to each shock in the five categories discussed above, and 95% confidence intervals are also included. Under the null hypothesis that dispersion is not affected by movements in aggregate output, the estimated responses should be distributed randomly around zero. This is overwhelmingly not the case. Instead, almost all of the responses are consistent with the interpretation that cross-sectional dispersion reacts endogenously to the business cycle in a particular direction: the dispersion measures increase after a negative aggregate shock (for example, monetary tightening, oil price increase) and they decrease after favorable developments (for example, government spending shock, increase in productivity). Only the responses to the Ramey news shock are at odds with this hypothesis, but they are highly insignificant.

In several cases, the estimated response is statistically significant at least at some lags at the 5% level, which is remarkable given the very limited amount of data for some of the shock series. Overall, credit supply disturbances, oil price shocks and the Romer and Romer monetary surprises seem to have a clear-cut effect on the dispersion measures. On the other hand, the effect of fiscal changes is estimated very imprecisely, so it remains insignificant in all cases.

It is interesting to look at the impulse response functions that analogous regressions imply in the other direction. One could argue that my aggregate shock series borrowed from the literature are mismeasured, and are capturing some effects of micro-level uncertainty. To the extent that cross-sectional dispersion is a valid proxy of idiosyncratic uncertainty, this could lead to spurious correlation between the aggregate shocks and dispersion. In this case, however, we would expect the statistical relationship to be present in both direction. That is, the aggregate shock series should respond to changes in micro uncertainty indicators. This prediction does not bear out in the data at all. The aggregate shock variables do not show any systematic pattern after innovations in dispersion. The estimated impulse responses often switch sign, and they are always insignificant.[4] This may increase our confidence that the co-movement between aggregate shock variables and dispersion is not driven by contamination by second moment shocks.

Using the shocks as external instruments

The uncertainty literature relies heavily on simple VARs when demonstrating the business cycle implications of second moment shocks. For example, Bachmann, Elstner and Sims (2013) estimate bivariate systems featuring a measure of firm-level uncertainty and a measure of aggregate economic activity. Statistical identification of uncertainty shocks is achieved through a triangular decomposition, which makes strong exclusion assumptions and depends on the ordering of variables. These adhoc assumptions are made because we lack an intuitive concept of exogenous changes in idiosyncratic uncertainty, one that is similar to our understanding of exogenous monetary policy, fiscal policy or oil price shocks.

I argue that despite the lack of exogenous variation in uncertainty, we can still use what we know about standard business cycle shocks to shed light on the relationship between dispersion-based proxies of uncertainty and output. I demonstrate the idea in a bivariate VAR setting similar to Bachmann, Elstner and Sims (2013). A bivariate system is a parsimonious way to model the joint dynamics of cross-sectional dispersion and activity. The frequency of the series in the VARs is quarterly, the VARs are estimated with 4 lags, and the activity variable, the logarithm of GDP, enters the system in levels.

I envisage two structural shocks driving the system: a level-shock and a second moment shock. I assume that a positive level-shock increases aggregate output, but its impact on cross-sectional dispersion is not restricted. Similarly, the second moment shock affects the dispersion of firm-level variables by assumption, but it does not necessarily have any effect on aggregate output. The goal is to estimate the unrestricted structural parameters using the results from the reduced-form VAR. To achieve this goal, I will use several shock series *jointly* as "external instruments" to identify the effect of level-shocks on dispersion.

The essence of this identification scheme is to find variables outside of the VAR that are correlated with the structural shock of interest, but are not correlated with any other structural shocks. A variable with these properties is called an external instrument. The structural shock is then identified – after a necessary normalization – from the correlation of the reduced-form VAR innovations and the external instruments.[5] The papers cited in section 1.2 construct exogenous components of specific first moment shocks directly, and treat these components as exogenous shocks. However, technically they are instrumental variables for the shocks: they are not the full shock series, but rather measure (typically with error) an exogenous component of the shock, so that the constructed series is correlated with the

[4] These insignificant impulse response functions are not reported to save space.

[5] The instrumental variables approach to structural VAR identification used in this exercise was originally presented in Stock and Watson (2008). The methodology was also developed independently in Mertens and Ravn (2013).

shock of interest but not with other shocks. This is exactly the requirement for a valid external instrument. The formal derivation of the GMM estimator and further details are given in Appendix A.2.

The crucial identifying assumption is that the aggregate shocks are valid instruments for the level-shock. Researchers who contributed to this literature clearly intended to capture exogenous first moment changes that have aggregate implications, so the question is how successful they were. In cases when the construction of the shock involved little or no model-based assumptions (for example, narrative evidence), it is likely that the series are not contaminated by some sort of second moment shock. In some other cases, the construction may also rely on theoretical assumptions, and one could argue that the authors unintentionally captured the effect of second moment shocks too. For example, Gilchrist and Zakrajsek (2012) try to purge credit spreads from the expected default component, which also reflects changes in volatility. However, their methodology necessarily relies on a theoretical model of the probability of default. If this model performs poorly, their measure of credit supply conditions will be contaminated by the effect of volatility shocks.

Nonetheless, this identifying scheme has several advantages over an adhoc Choleski ordering. First, there is no natural minimum delay assumption between dispersion and output, which is often invoked for the ordering of policy variables such as the short term interest rate. My methodology does not impose any restriction on timing, and it has a clear instrumental variables interpretation: the validity of the estimates hinges on the validity of the instruments. Second, one can check the sensitivity of the results to the set of shocks included in the estimation which mitigates somewhat the concerns about endogeneity. Furthermore, the GMM methodology allows for a formal over-identification test.

When estimating the baseline SVAR model, I use only one shock series from each of the five broad categories as external instruments. I have shown that individually almost all of them have a significant effect on dispersion, so ideally I would want to use all the twelve series together. Unfortunately, this is not possible because of data limitations. First, the time span of the various shock series is very different, and some of them are available only for a relatively short time period (see Table 1.1). To obtain meaningful estimates, the external instruments need to have sufficient overlap with each other. Second, too many moment conditions relative to the sample size can also lead to numerical instability (flat regions) in the optimization of the GMM objective. Considering these shortcomings of my dataset, I choose from each category the shock with the longest available time series, giving preference to shocks whose construction is not based on strong theoretical restrictions. The baseline results utilize the Hamilton (1996) oil price shock, the Romer and Romer (2004) monetary policy shock, the Basu, Fernald and Kimball (2006) technology shock, the Fisher and Peters (2010) fiscal shock, and the Gilchrist and Zakrajsek (2012) credit supply shock. I also carry out an extensive sensitivity analysis to the set of instruments in the Appendix.

The left-hand panels of Figure 1.11 plot the response of sales growth dispersion and stock return dispersion to an output level-shock in the baseline SVAR. The response of dispersion to a level-shock is highly statistically significant, and its magnitude is economically meaningful. Specifically, a 2 percentage points negative shock to GDP raises both dispersion measures by about one standard deviation on impact.[6] The right-hand panels show the same impulse responses, but after dropping the Gilchrist-Zakrajsek excess bond premium (GZ-spread) from the list of instruments. This is an important robustness check, because arguably this shock series is the most vulnerable to contamination by true idiosyncratic uncertainty shocks, as discussed above. It is reassuring that the results remain relatively unchanged after this modification: the response of stock return dispersion becomes smaller, but is still significantly different from zero for 8 quarters. Altogether, the impulse response functions confirm that business cycle fluctuations trigger significant movements in the dispersion of firm-level outcomes.

My identification strategy also allows me to recover the response of output to a second moment shock. The reason is that in a bivariate VAR, we only need to identify one parameter based on theoretical restrictions; the rest of the mapping between structural shocks and reduced-form innovations is pinned down by the covariance matrix of the VAR residuals. Making use of these extra moment conditions, I estimate the impulse response function of output to a second moment shock, and I also carry out a forecast error variance decomposition in the baseline SVARs. The impulse response functions in Figure 1.12 reveal that the response of GDP to a second moment shock which raises dispersion by one standard deviation is small and statistically not significant. The point estimates are actually positive, which is inconsistent with uncertainty-driven business cycle theories. Figure 1.13 suggests that the two structural shocks contribute roughly equally to the variance of cross-sectional dispersion over both short and longer horizons. In contrast, output movements are predominantly driven by level-shocks, and second moment shocks are responsible only for a moderate share of the forecast error variance. Note that the results are remarkably similar, both qualitatively and quantitatively, for the two dispersion measures.

Appendix A.1 documents that further variations in the set of instruments do not change the main results. It also presents formal over-identification tests in the baseline SVARs that fail to reject that the different external instruments identify the same causal relation between output and cross-sectional dispersion.

My empirical results are in stark contrast with the uncertainty interpretation of countercyclical dispersion. On the one hand, the cross-sectional spread of firm performance, both as perceived by the stock market and as measured by realized turnover, responds to changes in the level of output. On the other hand, this endogenous reaction seems to account entirely for the negative co-movement between dispersion and aggregate economic activity: innovations in dispersion have a positive and insignificant effect on output.

1.3 Ruling out mechanical explanations

A common interpretation of countercyclical cross-sectional dispersion is that dispersion measures proxy for idiosyncratic uncertainty, which is an independent driver of the business cycle.

[6]The standard deviation of sales growth dispersion and stock return dispersion are 0.03 and 0.015, respectively.

Using historical episodes and a series of formal tests, I have argued that the other direction of causality is more plausible: changes in dispersion are the result of fluctuations in output. But what is the mechanism? How does a first moment shock that pushes the economy into recession induce higher heterogeneity of firm-level outcomes? When trying to answer this question, one should take into account that countercyclical dispersion is a very robust feature of the micro data. It has been documented for different variables (quantities and prices) and in different markets (goods, labor and financial markets). This suggests that a fairly general mechanism is operating – one that works across a large class of models and is able to explain the dispersion of many variables.

One very general explanation is suggested by Jurado, Ludvigson and Ng (2013) in their critique of conventional uncertainty proxies. They point out that dispersion in firm-level variables can fluctuate over the business cycle merely because there is heterogeneity in the cyclicality of firms' business activity. Similarly, cross-sectional dispersion in individual stock returns can fluctuate without any change in uncertainty if there is heterogeneity in the loadings on common risk factors.[7] In other words, the high degree of symmetry of firms that most theoretical models maintain might be too strong an assumption when thinking about time-varying dispersion. In reality, different firms or different sectors of the economy may have different sensitivity to aggregate shocks. If this is the case, the heterogeneous impact of common shocks can lead to fluctuations in cross-firm dispersion. In this section, I study whether such mechanical effects can explain the causal link from aggregate output to cross-sectional dispersion in my data.

To organize the discussion, it is useful to think about firm-level variables as obeying a factor structure,

$$y_{it} = a_i + b_i f_t + \sigma_t \varepsilon_{it}, \tag{1.2}$$

where y_{it} can be any firm performance measure (for example, stock return), f_t is some aggregate factor capturing the state of the business cycle, and $\varepsilon_{it} \sim iid(0, 1)$ is an idiosyncratic shock (for example, productivity). This formulation allows for heterogeneity in trend growth (a_i) and cyclical sensitivity (b_i), and the variance of the idiosyncratic shock can be time-varying (σ_t). Let d_t denote the cross-sectional variance of y_{it} at time t. Then from (1.2) we can deduce the potential sources of time-varying dispersion:

$$d_t = var(a_i) + var(b_i) f_t^2 + cov(a_i, b_i) f_t + \sigma_t^2.$$

Note that theories in the "uncertainty" shock literature abstract from heterogeneity among firms, but postulate an exogenous process for σ_t. The mechanical explanations that I will test focus on the second and third terms containing f_t, since in the presence of cross-firm heterogeneities they can lead to correlation between dispersion and the business cycle.

[7] Jurado, Ludvigson and Ng (2013) do not study the causal link between aggregate output and dispersion. Instead, they want to provide direct econometric estimates of time-varying uncertainty. They mention heterogeneous loadings as a potential explanation for why their indicator is very poorly correlated with cross-sectional dispersion measures.

Consider the case of heterogeneous loadings on the aggregate factor, that is when $var(b_i) > 0$. This certainly implies that aggregate shocks can change cross-sectional dispersion. However, this is unlikely to drive my empirical findings for two reasons. First, the co-movement between output and dispersion is virtually unchanged if one removes the common component from firm-level variables before calculating dispersion. Specifically, I regress individual stock returns on the three factors of Fama and French (1992), and I regress firm-level sales growth on GDP growth and the change in the unemployment rate. In both cases, the cross-sectional dispersion of the regression residuals has the same cyclical properties as the raw dispersion series (see Appendix A.1). Second, notice that $var(b_i)$ is multiplied by the square of f_t, which implies that dispersion should increase equally for both unusually high and unusually low economic activity. Thus, one would need to appeal to a variant of (1.2) where the heterogeneity of "downside" sensitivities is much bigger than the heterogeneity of "upside" sensitivities. To test this hypothesis, I repeat my regressions allowing for different loadings for high (above average) and low (below average) values of the common factors. I cannot find evidence in the data that firm sensitivities are more dispersed in bad times than in good times.

Another mechanical explanation is negative correlation between trend growth and cyclical sensitivity across firms, that is $cov(a_i, b_i) < 0$. I refer to this possibility as the Abraham-Katz mechanism, since Abraham and Katz (1986) use exactly this reasoning to explain the positive correlation between the dispersion of employment growth rates across sectors and the aggregate unemployment rate. In my context, this story requires that slow growth firms have high cyclical sensitivity, while firms with higher trend growth should be relatively insensitive to the business cycle. I can test the Abraham-Katz mechanism using regressions of firm-level performance variables on measures of aggregate activity, and examining the relationship between the intercept and the sensitivity parameter.[8] Again, I do not find evidence that this could drive my results: if anything, $cov(a_i, b_i) > 0$ in my data (see Figure 1.14).

One could also argue that the sharp distinction between common shocks and idiosyncratic shocks in (1.2) is not the right way to think about firm-level variables over the business cycle, and that sector-specific shocks play a decisive role in aggregate fluctuations. For example, construction was hit disproportionately in the most recent recession, while information technology firms suffered the most in the 2001 downturn. When constructing cross-sectional dispersion measures, I pool together firms from all sectors. Consequently, sectoral shocks will change dispersion, even though within-sector dispersion might not move along the cycle. To address this concern, I re-calculate the dispersion time series after controlling for industry-quarter fixed effects in firms' sales growth and stock returns. This not only partials out the effect of sectoral shocks, but also deals with the possibility of different sectors having different responsiveness to aggregate shocks. Overall, the qualitative results are not affected. Figure 1.15 shows the residual dispersion for sales growth. The filtering does have an effect

[8] More specifically, I regress firm-level sales growth and stock returns on a constant and detrended GDP growth. I define trend growth as the intercept, and cyclical sensitivity as the coefficients on GDP growth. I also checked my results using the unemployment rate as the common factor.

on the level of dispersion, and there are some episodes when sectoral heterogeneity explains a nontrivial part of the increase in dispersion. However, the two series move together very closely and dispersion is still strongly countercyclical.

I conclude that the effect of the business cycle on the dispersion of firm-level outcomes cannot be attributed to the differential impact of aggregate shocks across firms or sectors. Imposing a factor structure on the data in order to explain the observed time variation in dispersion does not take us very far. In the end, the lion's share of the interesting business cycle variation always shows up in the residual. This suggests we need a mechanism that relates the aggregate state of the economy, f_t, to the idiosyncratic component in equation (1.2). In the next section, I present an equilibrium model whose reduced form can be summarized as

$$y_{it} = a + bf_t + \lambda(f_t)\varepsilon_{it},$$

with $\lambda(\cdot) > 0$ a decreasing function of f_t. As is clear from this expression, firms are perfectly symmetric in terms of their exposure to an aggregate (TFP) shock. However, their responsiveness to an idiosyncratic shock, which has constant volatility σ, varies according to the aggregate state.

1.4 An attention-based model of dispersion

This section develops a model to explain time-varying dispersion without second moment shocks.

Motivation and main intuition

In the recent global financial crisis and recession, many observers made statements suggesting that attention had become much more focused on various forms of heterogeneity. For example, the famous comment by Warren Buffet, *"You don't know who's naked until you drain the swimming pool"*,[9] suggests that investors learn more about firm-level characteristics when the aggregate economy is suffering. In the academic literature, Loh and Stulz (2014) show that stock market participants indeed behave in a way consistent with this view. The authors examine a large sample of analyst output on individual stocks, and find that analysts work harder in bad times, and their revisions to earnings forecasts and stock recommendations have a much more influential stock-price impact during bad times. Tyler Cowen in a New York Times column points to a similar mechanism in relation to worker experiences in the Great Recession: *"Before the crisis, for example, business executives and owners didn't always know who their worst workers were...So long as sales were brisk, it was easier to let matters lie. But when money ran out, the axes fell...and the crash led to this information being revealed."* (Cowen, 2014) In the context of Europe's woes, the remarkable uniformity of risk measures (for example, long term interest rates) before the debt crisis

[9] Cited by Mark Gertler in Price (2013).

also has an interpretation that markets did not process carefully the information about the inherent heterogeneities of these countries. It took a large negative shock to call attention to these differences.

In the model, I use countercyclical information acquisition to generate the observed time-variation in cross-firm dispersion. Heterogeneity is introduced through idiosyncratic shocks which have a constant volatility. However, gathering information about these idiosyncratic shocks is costly, so firms may rationally choose to ignore them when making decisions. Uninformed firms will tune their decisions to the aggregate state of the economy, which results in more homogeneous firm-level outcomes. The setup provides a unifying interpretation for the countercyclical variation in the dispersion of four micro-level variables: (i) price changes, (ii) firm sales growth, (iii) firm real value-added and (iv) stock returns.

In the model presented below, I use a very simple mechanism to rationalize countercyclical information acquisition: firms have to sacrifice inputs to overcome the informational friction, and the cost of these inputs is low in bad times. However, in Appendix A.3 I discuss other frictions that can realistically lead to the same effect, and I present an illustrative partial equilibrium model where the threat of costly external financing induces more attention in bad times.

Preferences and technology

The model features a representative household who maximizes

$$E_t \sum_{s=0}^{\infty} \delta^s \left[\frac{C_{t+s}^{1-\theta}}{1-\theta} - \frac{\chi}{\gamma} L_{t+s}^{\gamma} \right]; \quad \theta > 1, \gamma > 1, \chi > 0$$

subject to the period budget constraint

$$B_{t+s} + P_{t+s} C_{t+s} = w_{t+s} L_{t+s} + R_{t+s-1} B_{t+s-1} + D_{t+s}.$$

C_t is a consumption basket with an aggregate price index P_t, and w_t is the nominal wage. B_t denotes one-period bond holdings at the end of time t, which pays the gross risk free nominal interest rate R_t in the next period. Firms redistribute all profits as lump sum dividends D_t.[10]

The consumption aggregator is defined as

$$C_t = \left[\int_0^1 c_{it}^{\rho} di \right]^{\frac{1}{\rho}},$$

where $i \in [0, 1]$ indexes the goods, $0 < \rho < 1$ and $\sigma = 1/(1-\rho)$ is the elasticity of substitution.

[10] By assuming lump-sum dividends, I abstract from the stock market. However, using the household's stochastic discount factor I can value the firm, and calculate hypothetical stock returns. I will do so when investigating the relationship between fluctuations in output and stock return dispersion.

Optimal consumption allocation yields the good-specific demand

$$c_{it} = C_t \left(\frac{p_{it}}{P_t} \right)^{-\sigma}, \tag{1.3}$$

where $P_t = \left[\int_0^1 p_{it}^{1-\sigma} di \right]^{1/(1-\sigma)}$ is the aggregate price index. Let $R_t = P_t C_t$ denote aggregate total expenditure in the economy.

The supply side of the economy is characterized by a continuum of monopolistically competitive firms producing differentiated varieties. The production technology is given by

$$y_{it} = A_t z_{it} l_{it}, \tag{1.4}$$

where A_t is an aggregate productivity shock that follows an AR(1) process in logs, and z_{it} is an i.i.d. idiosyncratic productivity shock distributed as $\log z_{it} \sim N(0.5\alpha^{-1}, \alpha^{-1})$.[11] Note that the volatility of firm fundamentals is constant over time. In order to focus on the effect of information acquisition on cross-sectional dispersion, I set up the environment so that firms carry out static profit maximization.

Information structure

Every period firms set their prices subject to an information constraint: they observe all aggregate variables, but they cannot directly observe the idiosyncratic productivity shock. If they want to learn about z_{it}, they have to devote labor resources to information acquisition. The result of information gathering efforts is a noisy signal of the true idiosyncratic productivity: $s_{it} = \log z_{it} + \beta_t^{-0.5} \varepsilon_{it}$ where ε_{it} is i.i.d. standard normal across firms and time. Firms are able to choose β_t, the precision of the signal. The cost of information depends on this precision, and it is given by $G(\beta)$, expressed in labor units, where $G(0) = 0$, $G'(\beta) \geq 0$, $G'(0) = 0$, and $G''(\beta) > 0$. The more work the firm puts into information gathering, the more precise information it will have.[12]

I capture micro-level heterogeneity of firm environments as different idiosyncratic productivity levels. Arguably, if a firm is in operation for a while, it can observe how much input it needs to produce a unit of output. One interpretation is that the firm has to set its price *before* starting production, and once that price is announced, it has to satisfy any demand for its product. One can think of $G(\beta)$ as the amount of labor that is needed to test the production process before actual production starts; the more careful the testing (higher β), the more labor is needed. Another, perhaps more intuitive, interpretation is based on the fact that productivity differences may reflect not only cost differences, but also differences

[11] The mean of the idiosyncratic productivity distribution is chosen to simplify algebraic expressions, and has no other implication in the model.

[12] We can imagine that information comes in small packages, each containing a piece of independent data about z_{it}. Finding new packages requires more and more effort. In this environment, choosing β amounts to deciding the number of packages to collect, and the convex information cost in labor units arises naturally.

in consumer valuations of the good. For a firm operating in a constantly evolving market environment, it might be hard to know the exact demand curve it faces. In this case, $G(\beta)$ can be interpreted as the labor input of the marketing research project that helps uncover the consumer valuation of the firm's product. The general idea is that with more effort the firm can achieve a more precise understanding of its business environment, and it can fine-tune its decisions.[13]

The assumption that the information acquisition cost is measured in labor units, as opposed to units of the consumption good, formalizes the idea that the technology of information acquisition is distinct from the technology of producing goods, and that it is more stable. Whether the economy is in a recession or a boom, the amount of labor needed to produce micro-level information is the same. I view this assumption as more reasonable than the other extreme when the information cost is fixed in consumption units. Modeling market access costs in labor units has a long tradition in trade and in dynamic industry equilibrium models (for example, Hopenhayn, 1992; Melitz, 2003). In that literature, these costs are often justified as the necessary labor input or time to learn about market-specific regulations, to build relationships with new retailers, or to conduct marketing studies about demand in the new market. My assumption is very similar in spirit, with a minor difference: the information firms acquire is noisy, and in principle a firm can start production without collecting any information, relying only on the common prior distribution of productivities.

Profit maximization

Within each period, firm managers have to solve a two-stage profit maximization problem. First, they pick the signal precision, that is they decide how much attention to pay to idiosyncratic conditions. Next, they set their price based on the acquired information embodied in the signal. Relegating most of the tedious algebra to Appendix A.4, I proceed to solve this problem backwards.

Consider the second stage, that is the price-setting decision for given signal precision. A firm with realized signal value s maximizes its expected profit subject to the demand constraint. Let $v(s)$ denote the value function of this optimization:

$$v(s) = \max_p E_s \left[\left(p - \frac{w}{Az} \right) \left(\frac{p}{P} \right)^{-\sigma} Y \right],$$

where E_s is short for the conditional expectation operator $E[.|s]$.[14] This yields the usual

[13] With CES demand and monopolistic competition, differences in productivity across firms have identical effects on equilibrium revenue and profits as differences in consumer tastes. I chose the productivity interpretation because it generates price dispersion, which would not be the case with firm-specific demand shifters. This is only a technical curiosity of CES demand, rather than anything substantial about the real world.

[14] Since the price-setting problem is static, and s is the only relevant firm-level variable, I drop the time and firm subscripts. Later they will be reintroduced when I study the time series properties of cross-sectional dispersion. I also impose the goods market clearing condition $C = Y$.

markup formula for the optimal price

$$p(s) = \frac{w}{\rho A \bar{z}(s)}, \tag{1.5}$$

where

$$\bar{z}(s) = \exp\left[\frac{\beta}{\alpha+\beta}s\right] \tag{1.6}$$

can be interpreted as the firm's best estimate for its productivity level based on the information it had acquired earlier. Importantly, this estimate is shaped by β, since the interpretation of s is contingent on the the precision of the signal.

Plugging back the optimal price, we can write the second-stage expected profit (without the information cost) of a firm with signal s as

$$v(s) = \frac{R}{\sigma}\left(\rho A \bar{z}(s)\frac{P}{w}\right)^{\sigma-1}. \tag{1.7}$$

Consider now the first stage when the firm decides how much effort to put into information gathering. It has to compare the expected profit gain from more precise knowledge of z to the cost of acquiring information. Taking expectations of (1.7),[15] the ex ante expected profit of a firm that chooses precision β is

$$\pi^e(\beta) = \frac{R}{\sigma}\left(\rho A \tilde{z}\left(\beta\right)\frac{P}{w}\right)^{\sigma-1} \tag{1.8}$$

where

$$\tilde{z}(\beta) = \exp\left[\frac{1}{2}\frac{\beta}{\alpha+\beta}\sigma\alpha^{-1}\right]. \tag{1.9}$$

Expected profit is increasing in signal precision, with the magnitude of the effect depending on the elasticity of substitution. The higher is σ, that is, the closer we are to perfect competition, the more valuable is information. The intuition is easy to see – when goods are highly substitutable, mispricing due to imprecise perception of productivity is particularly costly. The following proposition establishes that under a mild condition the firm's two-stage profit maximization problem has a unique solution that can be characterized by a first order condition.

Proposition 1 *Given aggregate variables (P, Y, w), the firm's ex ante expected profit in (1.8) is strictly increasing in signal precision, that is $d\pi^e(\beta)/d\beta > 0$. If $\frac{1}{4}(\sigma-1)\sigma - \alpha < 0$, then the expected profit function is also strictly concave on the whole domain of $\pi^e(\cdot)$, that is $d^2\pi^e(\beta)/d\beta^2 < 0$. Under this assumption, the firm's information acquisition problem*

$$\max_{\beta} \pi^e\left(\beta\right) - wG\left(\beta\right)$$

[15] As an abuse of notation, I use s and z to refer to random variables from an ex ante perspective and also to denote the realized values of these random variables for a given firm.

has a unique solution β^* which can be fully characterized by

$$\pi^{e\prime}(\beta^*) = wG'(\beta^*). \tag{1.10}$$

Proof. See Appendix A.4. ■

The technical condition $\frac{1}{4}(\sigma-1)\sigma-\alpha<0$ ensures that the information acquisition problem is globally concave. Otherwise, the marginal value of information is increasing for low levels of β (in particular, for $\beta<\frac{1}{4}(\sigma-1)\sigma-\alpha$). It is interesting to note that this peculiarity is not specific to my setting, but it is merely the manifestation of an unresolved general problem in the theory of information demand. In a well-known paper, Radner and Stiglitz (1984) show that under fairly general smoothness assumptions, the value of information exhibits increasing marginal returns over some range. Trying to understand and overcome this paradox is an active research agenda in information economics. For example, Moscarini and Smith (2002) show that if we measure the quantity of information by the number of independent observations from an experiment, then the marginal value of information eventually falls as the number of observations increases. This result applies to my setting since a high precision signal is equivalent to observing a large number of independent low precision signals. In numerical experimentations, I never find the lack of global concavity a problem; nevertheless, I make sure that in my calibration this technical constraint is satisfied.

Aggregation and equilibrium

Since all firms are ex ante identical, I look for a symmetric Nash equilibrium where all firms acquire the same amount of information. I solve the model by backward induction. First, for a common signal precision β, I derive the aggregate variables that the individual firm takes as given. Then I find the equilibrium information acquisition strategy β^{EQ} as the fixed point of the firm's best response function.

I choose labor as the numeraire, and set $w=1$. Plugging in the optimal price from (1.5) into the aggregate price index we obtain

$$P = \frac{1}{\rho A\tilde{z}(\beta)}, \tag{1.11}$$

where $\tilde{z}(\beta)$ is defined in (1.9). Using this expression together with the production technology, the demand function, and the price-setting rule, we can write aggregate labor demand as

$$L^D = \frac{Y}{A\tilde{z}(\beta)} + G(\beta), \tag{1.12}$$

where the first term is the amount of labor used in the goods production, while the second term is labor used in the production of information.

Combining labor demand with the representative household's labor supply schedule

$$L^S = \chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}} \left(\frac{1}{P}\right)^{\frac{1}{\gamma-1}},$$

we can determine equilibrium output for a given β. Although no closed-form solution can be given, it is easy to establish that there is a unique intersection of the labor supply and labor demand curves.

At this point, we can express all the aggregate variables (L, Y, P, R) as functions of the common signal precision. Now let us assume that in the symmetric Nash equilibrium each firm acquires β^{EQ} amount of information. Substituting the aggregate variables into the firm's first order condition in (1.10) and imposing the fixed point property ($\beta^* = \beta^{EQ}$), we obtain a simple characterization of the equilibrium signal precision:

$$\frac{1}{2}(\sigma-1)R\left(\beta^{EQ}\right)\left(\alpha+\beta^{EQ}\right)^{-2} = G'\left(\beta^{EQ}\right). \tag{1.13}$$

The notation $R\left(\beta^{EQ}\right)$ makes it explicit that aggregate expenditures depend on β^{EQ}.

The following proposition completes the solution of the model.

Proposition 2 *There is a unique β^{EQ} that satisfies (1.13). If $\frac{1}{4}(\sigma-1)\sigma-\alpha<0$, then the first-order condition is sufficient to characterize the best response correspondence of the firm; thus β^{EQ} is the unique Nash equilibrium of the model.*

Proof. Appendix A.4. ■

The intuition for the existence of a unique symmetric equilibrium is the following. The labor demand condition in (1.12) reveals that imperfect information about idiosyncratic fundamentals leads to a misallocation of resources among firms, reducing aggregate productivity and output. The size of this effect depends on the quality of information at the firm level. For $\beta = 0$, aggregate productivity is simply A, so the economy loses all the potential gains from allocating more resources to more productive firms. Higher β moves aggregate productivity closer to the frictionless outcome. However, from the individual firm's perspective a more efficient aggregate economy is not desirable. If all other firms are well-informed, the market will be more competitive – the aggregate price level in (1.11) will be low. This reduces the expected profit of an individual firm, and makes the return to information acquisition small. Hence, there is strategic substitutability in information acquisition – if everybody else learns a lot, then each firm wants to learn only a little – which leads to a unique interior equilibrium.

1.5 Information Acquisition, Dispersion and the Aggregate State

The equilibrium of this simple endogenous information model is consistent with the stylized facts about countercyclical dispersion. The argument has two pieces. First, higher aggregate productivity leads to less information acquisition in equilibrium. The intuition is that using labor in the production of goods becomes more profitable relative to using it in information gathering. Second, less micro-level information leads to lower cross-sectional dispersion. Intuitively, when firms are uninformed about their own idiosyncratic shock, they base their decisions on the economy-wide average, which leads to convergence of outcomes. Notice that the second step of the argument is independent from the first piece. Any mechanism that encourages more attention in bad times will lead to countercyclical dispersion. The mechanism put forward here is just one example of possible explanations. Appendix A.3 presents an alternative intuition for why attention might vary over the business cycle.

In this setting, the precision of micro-level information affects the pricing decision, and then price dispersion shapes the dispersion of all the other firm-level variables (output, sales, profits, stock returns). Hence, it is enough to show that price dispersion is countercyclical.[16] Below I demonstrate analytically that this is the case, and then provide numerical simulations to show that the magnitude of the effect is in line with my empirical estimates.

Analytical results

The following comparative statics result shows that information acquisition is countercyclical.

Proposition 3 *The equilibrium β^{EQ} characterized in Proposition 2 is a strictly decreasing function of aggregate productivity, that is $d\beta^{EQ}/dA < 0$.*

Proof. Apply the implicit function theorem on equilibrium conditions. See Appendix A.4. ■

The assumption $\theta > 1$ in the household's utility function plays an important role in the proof. To interpret this assumption, notice that the θ parameter governs the strength of the income effect on labor supply. I need a sufficiently strong income effect so that the profit function of the firm grows more slowly than the wage in A. With logarithmic utility an increase in A leads to a proportional increase in both profits and the real wage. Given that the information cost is measured in labor units, this proportional change implies that the relative value of information is unaffected by aggregate productivity. In addition to being necessary for the mechanism to work, the imposed restriction on θ is completely in line with empirical estimates. Both risk aversion and elasticity of intertemporal substitution estimates establish 1 as a lower bound on this parameter. Admittedly, this implies a counterfactually procyclical real wage which is an undesirable feature of most flexible wage business cycle models.

[16] Existing business cycle models of micro uncertainty usually rely on heterogeneous firms that produce a homogeneous good with decreasing returns to scale technology (e.g., Gilchrist, Sim and Zakrajsek (2014), Bloom et al. (2012), Bachmann and Bayer (2013)). This makes it possible to pin down the firm size distribution, but also implies that they cannot address price dispersion. My model is able to capture the cross-firm dispersion of both prices and production, which I view as an advantage of the monopolistic competition setting.

The model implies a direct link between equilibrium attention and realized cross-firm dispersion. The following proposition derives a closed-form expression for the cross-sectional dispersion of percent price changes as a function of information acquisition.

Proposition 4 *The cross-sectional standard deviation of firm-level log price changes between $t-1$ and t is given by*

$$\sigma_{\Delta p,t} \equiv \sqrt{Var_i\left(\log \frac{p_{it}}{p_{i,t-1}}\right)} = \sqrt{\left[\frac{\beta_{t-1}}{\alpha+\beta_{t-1}} + \frac{\beta_t}{\alpha+\beta_t}\right]\alpha^{-1}}$$

where $Var_i(\cdot)$ indicates that the variance is taken across firms.[17]

Proof. Appendix A.4. ■

Since total output Y is driven by the persistent stochastic process of total factor productivity, Propositions 3 and 4 imply the following corollary.

Corollary 5 *Times of high (low) output tend to be associated with low (high) price change dispersion.*

Numerical illustration

This static model is designed to demonstrate the effect of endogenous information acquisition on cross-sectional dispersion. While the model is very simple, I show in a rough calibration exercise that it is able to generate simulated dispersion series and business cycle correlations that are close to their empirical counterparts. Needless to say, this should be viewed only as an illustration that the proposed mechanism has the potential to deliver meaningful quantitative results. Before presenting the results, I introduce a crude way to think about stock return dispersion within the model.

Calculating Stock Returns

Shares in firms are not traded directly. However, one can calculate the price that shares would have using the stochastic discount factor of the representative household. Let q_{it} denote the ex-dividend price of the stock of firm i. That is, the owner of the stock at time t is entitled to the profit stream of the firm from $t+1$ on. Then the price of the stock satisfies

$$q_{it} = E_t\left[\beta\frac{\Lambda_{t+1}}{\Lambda_t}\left\{\pi_{i,t+1}\left(z_{i,t+1}, s_{i,t+1}; \beta_{t+1}\right) + q_{i,t+1}\right\}\right],$$

where $\Lambda_t = Y_t^{-\theta}/P_t$. Observe that q_{it} will fluctuate over time due to the aggregate productivity shocks. Yet, in each period this price will be the same for all firms because idiosyncratic

[17] For readability, I drop the EQ superscript.

productivity and signal noise is i.i.d. Hence, if we define realized stock returns as

$$r_{it} = \frac{\pi_{it}(z_{it}, s_{it}; \beta_t) + q_{it}}{q_{i,t-1}} - 1, \tag{1.14}$$

then all the cross-sectional variation in returns will come from the dividend term.

This formulation makes the simulation of stock return dispersion straightforward. We can augment the system of equilibrium conditions with the pricing equation for a representative stock, and obtain the common stock price as a function of the aggregate state. Then we simulate realized profits for a large number of firms, and use (1.14) to calculate stock returns.

Calibration and Results

The model has at least two components that do not have evident empirical counterparts. Under my assumptions, the true dispersion of the idiosyncratic fundamental is never observed. Realized dispersion is not a valid measure, because it is shaped by both the underlying heterogeneity and the attention allocated to it. Likewise, there is no obvious real life analogue to the information cost function, $G(\cdot)$. Despite these difficulties, I try to impose some discipline on the calibration exercise through the following steps.

First, I fix the preference parameters at values which are well within the standard range used in the literature. Since my data is quarterly, I use $\delta = 0.99$ as the household's discount factor. I set the elasticity of substitution $\sigma = 4$, the risk aversion parameter $\theta = 4$, and the disutility of labor $\gamma = 3.5$, which implies a Frisch labor supply elasticity of 0.4. Second, I parameterize the aggregate shock process to match the persistence and volatility of HP-filtered log real GDP. This implies that aggregate productivity follows $\log A_t = 0.9 \log A_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim N(0, 0.015^2)$. Third, I calibrate the spread of idiosyncratic productivities by taking the estimated maximum dispersion of establishment-level manufacturing productivity that Bloom et al. (2012) report over 1968-2010. My rationale is the following: In my model, the largest possible realized dispersion is equal to the dispersion of fundamentals, and it is approached only if firms acquire close to perfect information. This should happen in very bad states of the world, when the economy is in a deep recession. Indeed, the peak dispersion according to Bloom et al. (2012) was achieved in the Great Recession. This logic leads me to set $\alpha = 4$ in the model.

Finally, I assume a simple linear functional form for G, and choose the marginal cost of signal precision to match average sales growth dispersion.[18] The average interquartile range of firm-level sales growth in my dataset is 0.22 (see Figure 1.1). It turns out that hitting this target is possible with $G(\beta) = 0.17\beta$. To interpret this information cost function, I calculated the share of labor hours devoted to information gathering in the simulated series (see Figure 1.16). During the biggest "boom" in the simulation, GDP is 5.1% above trend

[18] Strictly speaking, my assumptions in the model section do not allow for a linear cost function. Those assumptions guarantee the existence and uniqueness of equilibrium under general conditions, but they are by no means necessary.

and only 0.6% of total labor is used in information acquisition. In the deepest "recession" GDP is 4.0% below trend, but the share of hours devoted to information gathering quadruples to 2.4%. Although the proportional change is substantial, neither of these numbers seems to be unrealistically large.

I simulate 10,000 firms for 200 periods.[19] Assuming quarterly data, this loosely corresponds to the past 50 years for which I have data on sales growth and stock return dispersion. Figure 1.16 depicts some results from a typical simulation. The top two panels show the time series of the interquartile range of sales growth and realized stock returns together with the output gap, which is defined as the percent deviation from the non-stochastic steady state. As I showed analytically, a higher positive output gap coincides with lower cross-sectional dispersion. Note that although the calibration only tries to match average sales growth dispersion, the range and volatility of the simulated series are also remarkably close to the data.[20] The model also correctly predicts that stock return dispersion is smaller than sales growth dispersion, but the simulated range of the series is four times smaller than in the data.[21] This reflects the fact that firm profits (the dividend term) is only a small part of realized stock returns, and this is the only source of stock return dispersion in the model. The bottom panels present the correlation between year-on-year output growth and cross-sectional dispersion, which is another statistics often calculated in the literature. Again, the numbers are not out of line. In my data the correlation coefficient between quarterly sales growth dispersion and GDP growth is -0.35, while the simulated value is -0.32. The corresponding figures for stock return dispersion are -0.4 and -0.22.

Overall, the numerical results suggest that time-varying attention is able to generate quantitatively important movements in cross-sectional dispersion without any second moment shocks.

1.6 Conclusion

In this paper, I revisit some of the empirical evidence underlying the literature on "uncertainty" driven business cycles that has attracted much attention both in academic research and among policy-makers. I find that dispersion-based indicators of micro-volatility are unlikely to be valid proxies for exogenous second moment shocks. Instead, cross-sectional dispersion measures fluctuate endogenously over the business cycle. I propose a theory of cross-firm dispersion based on time-varying rational inattention.

Authors in the "uncertainty literature" often point to time-varying cross-sectional dispersion of firm-level variables, labeled "micro volatility". Micro-volatility is countercyclical; that is, recessions are more disparate across firms than booms. This has been interpreted

[19] I use a second-order perturbation in Dynare to solve for the rational expectations equilibrium value of the stock price.

[20] In my data, the time series standard deviation of sales growth dispersion is 0.033. The simulated counterpart is 0.030.

[21] The coefficient of variation is 0.22 in the data and 0.05 in the simulation.

as a result of a wider distribution of shocks hitting individual firms, which can have negative effects on aggregate output in the presence of real or financial frictions. I critique this argument in three steps.

First, I show that the direction of causality seems to run the other way: fluctuations in aggregate output induce movements in cross-sectional dispersion. To achieve causal identification, I isolate exogenous changes in aggregate output that were triggered by known first-moment shocks. Applying the recent econometric technique of "external instruments", I show that these exogenous movements in output are associated with significant changes in firm-level sales growth and stock return dispersion. Furthermore, innovations in dispersion – true second moment shocks – have an insignificant and wrong-signed impact on output. These results are inconsistent with theories based on uncertainty shocks.

Second, I examine whether mechanical effects, originating from firms' heterogeneous sensitivities to common shocks, can explain my findings. I identify the conditions under which it could be the case, and show that those restrictions do not bear out in the data. This motivates me to lay out a new theory that can explain countercyclical dispersion without invoking uncertainty shocks or asymmetric responsiveness of firms.

Third, I develop a simple model in which information about idiosyncratic fundamentals is imperfect, but firms can overcome this friction by sacrificing resources. I prefer to interpret this as the opportunity cost of finite attention or cognitive capacity, but it could also represent the actual cost of gathering information. When firms remain uninformed about their fundamentals, they rely more on the aggregate state of the economy, which leads to similar outcomes. If firms choose to pay more attention to their idiosyncratic business conditions in bad times, then recessions naturally emerge as times of high micro-volatility. In the model this happens because the cost of inputs used in information acquisition is low when aggregate productivity is low. However, in an appendix I briefly illustrate how other mechanisms, such as higher probability of disastrous outcomes, can also induce more attention in recessions.

I see two broad messages emerging from my analysis – one for policy-making and one for future theoretical work. First, more empirical research needs to be done to understand the role of uncertainty shocks in shaping business cycles. The evidence, at least for micro-level uncertainty shocks, is probably weaker than suggested by the salient place of uncertainty in public discourse. Similarly, if policy-makers want to mitigate fluctuations in micro-volatility, then it is important to understand what leads to this volatility. Second, endogenous information choice seems a promising modeling approach to explain economic phenomena when the degree of observed heterogeneity is time-varying. For example, countercyclical wage dispersion could be the result of employers paying more attention to idiosyncratic worker productivity after big negative shocks.

There are many directions in which this research could be extended in the future. On the empirical side, the most obvious extension is a more direct way of testing the model mechanism. I discussed some indirect and anecdotal evidence about heightened attention to heterogeneity after negative aggregate shocks. Unfortunately, directly measuring information flows is very challenging. In my view, financial markets may provide a good laboratory for this task. In future work, I intend to carry out a high frequency analysis of asset price dispersion in a short window around aggregate news announcements (for example, the employment report). If below-expectations reports are followed by more heterogeneous movements than above-expectations reports, then we can plausibly conclude that market participants have increased their attention to idiosyncratic characteristics. On the modeling side, embedding endogenous attention into dynamic business cycle models with other frictions (sticky prices, financial constraints, etc.) would be necessary to better understand the timing of information acquisition.

Table 1.1: External instruments used in the SVAR analysis

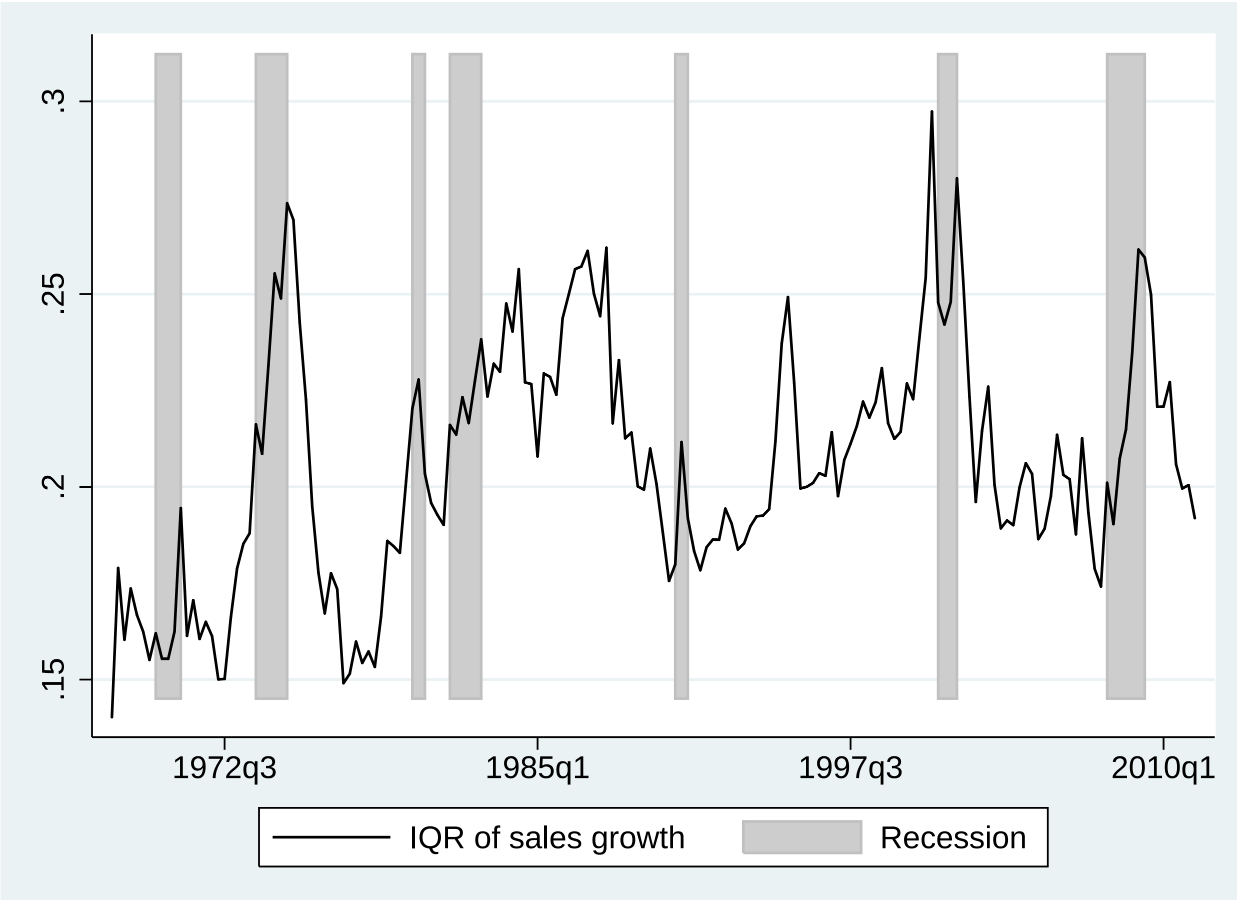
Name	Sample period	Source paper	Data source	Notes
		Oil		
OIL_HAM	1960Q1-2009Q4	Hamilton (1996)	BLS (oil PPI)	Quarterly average
OIL_KILIAN	1971Q1-2004Q3	Kilian (2008)	Lutz Killian website	
		Monetary policy		
MP_RR	1969Q1-2004Q4	Romer and Romer (2004)	Yuriy Gorodnichenko	Quarterly sum
MP_GSS	1990Q1-2004Q4	Gurkaynak, Sack, and Swanson (2005)	Eric Swanson website	Quarterly sum
MP_VAR	1965Q1-2007Q3	Christiano et al. (1999)	author's calculation	
		Productivity		
TECH_BFK	1960Q1-2012Q4	Fernald (2012)	John Fernald website	Series dtfp_util
TECH_GALI	1960Q1-2012Q3	Gali (1999)	author's calculation	
		Fiscal		
FISC_RR	1960Q1-2007Q4	Romer and Romer (2010)	David Romer website	Series EXOGENRATIO
FISC_RAMEY	1960Q1-2008Q4	Ramey (2011)	Valerie Ramey website	
FISC_EXRET	1960Q1-2008Q4	Fisher and Perters (2010)	Mark Watson website	
		Credit supply		
CR_BCDZ_BL	1992Q1-2010Q4	Bassett, Chosak, Driscoll, and Zakrajsek (2011)	Mark Watson website	
CR_GZ_EBP	1973Q1-2010Q3	Gilchrist and Zakrajsek (2012)	Mark Watson website	

Table 1.2: Dispersion lags the cycle (Granger Causality Tests)

Null Hypothesis	**chi2**	**df**	**Prob>chi2**
Sales growth disp. does NOT Granger cause GDP	1.84	4	0.766
GDP does NOT Granger cause Sales growth disp.	14.90	4	0.005
Stock return disp. does NOT Granger cause GDP	7.44	4	0.114
GDP does NOT Granger cause Stock return disp.	11.07	4	0.026

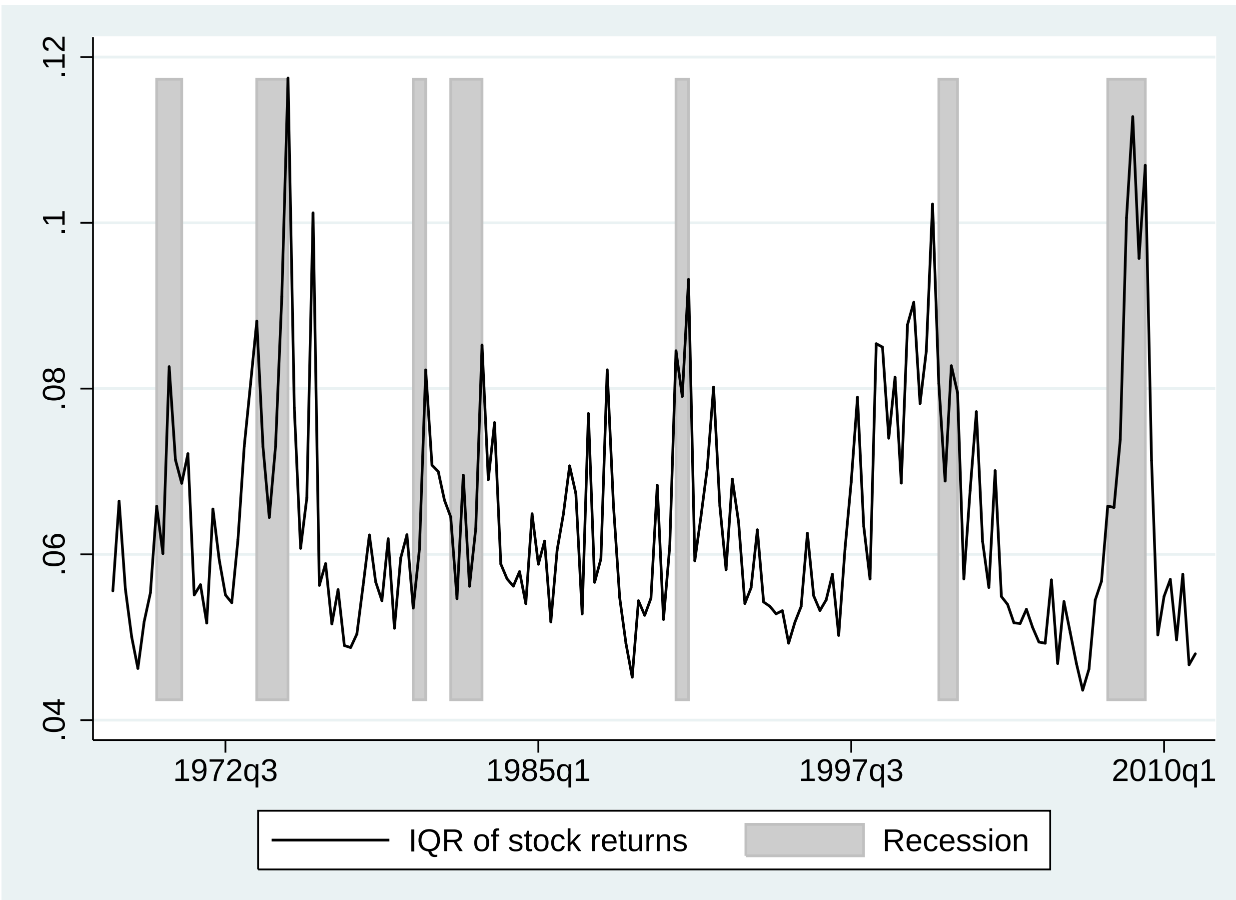
Notes: The test equations are estimated on quarterly data from 1968Q1 through 2012Q4, and include 4 lags of the dispersion measures and the output gap. Dispersion measures are defined in the text. The output gap is the log-deviation of real GDP from potential output as estimated by CBO.

Figure 1.1: Sales growth dispersion and NBER recessions



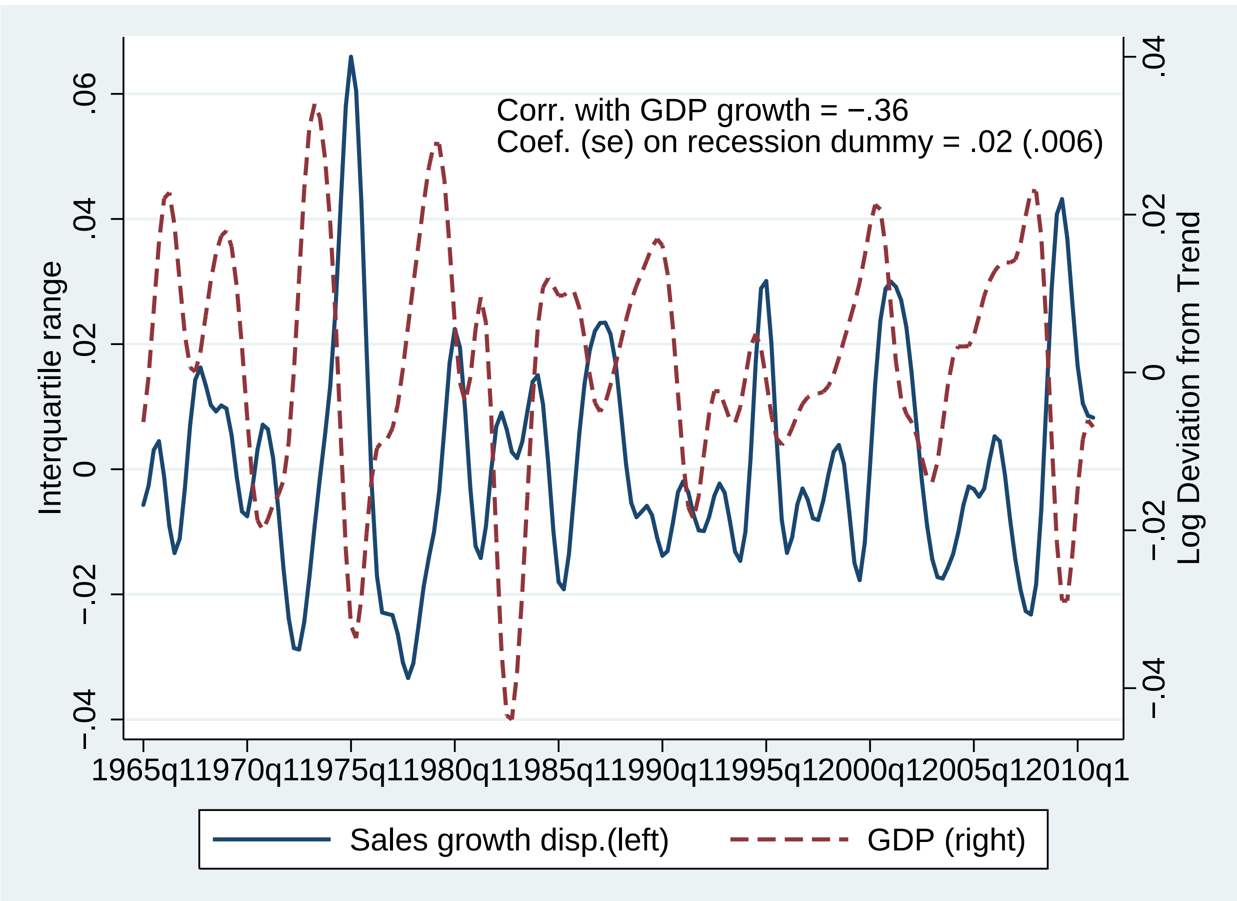
Notes: Sales growth dispersion in a given quarter is defined as the interquartile range of firm-level year-on-year sales growth. The sample includes Compustat firms with 25+ years of data.

Figure 1.2: Stock return dispersion and NBER recessions



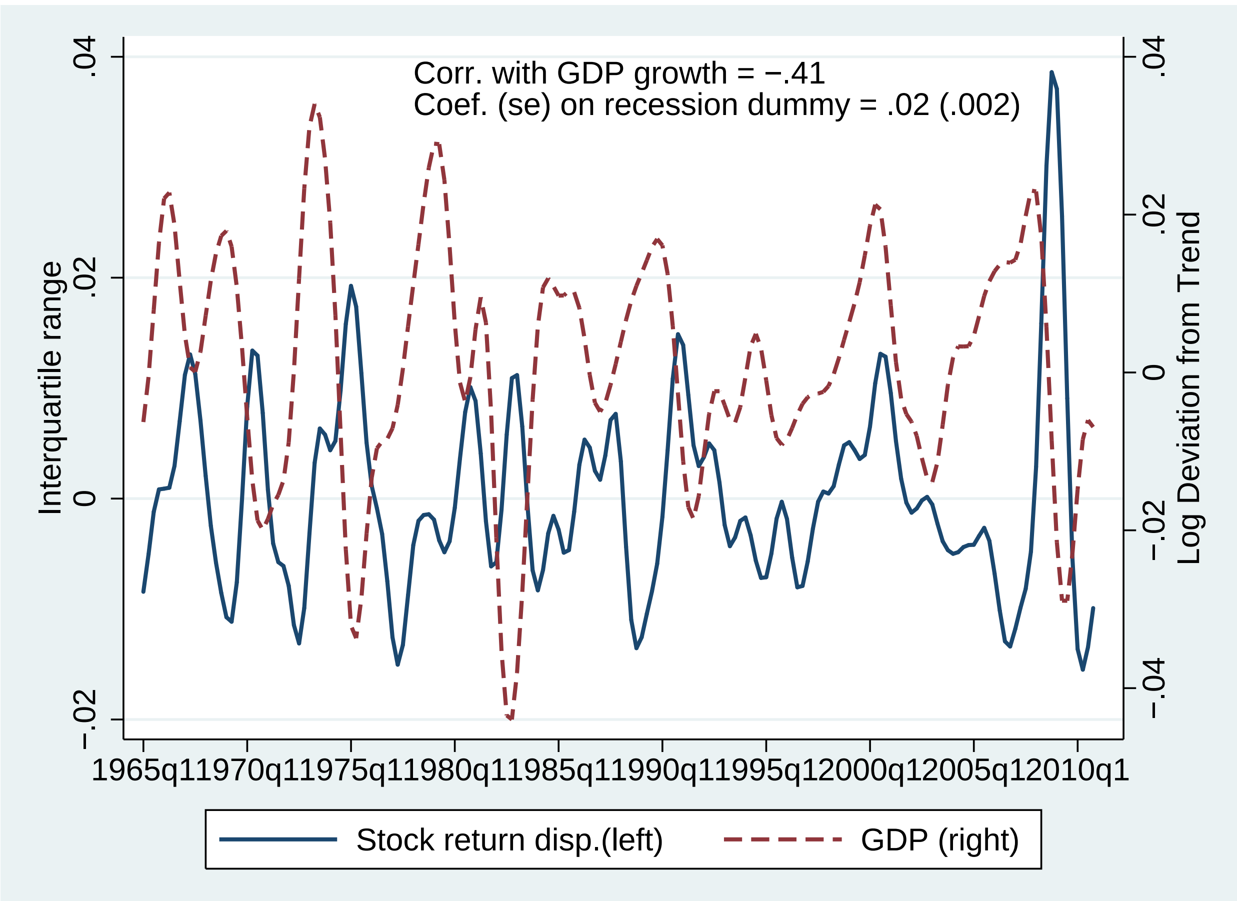
Notes: Stock return dispersion is defined as the interquartile range of firm-level stock returns in a given quarter. The sample includes CRSP firms with 25+ years of data. In the few instances when a firm has multiple traded securities, I use the one with the longest life span in the sample.

Figure 1.3: Sales growth dispersion over the business cycle



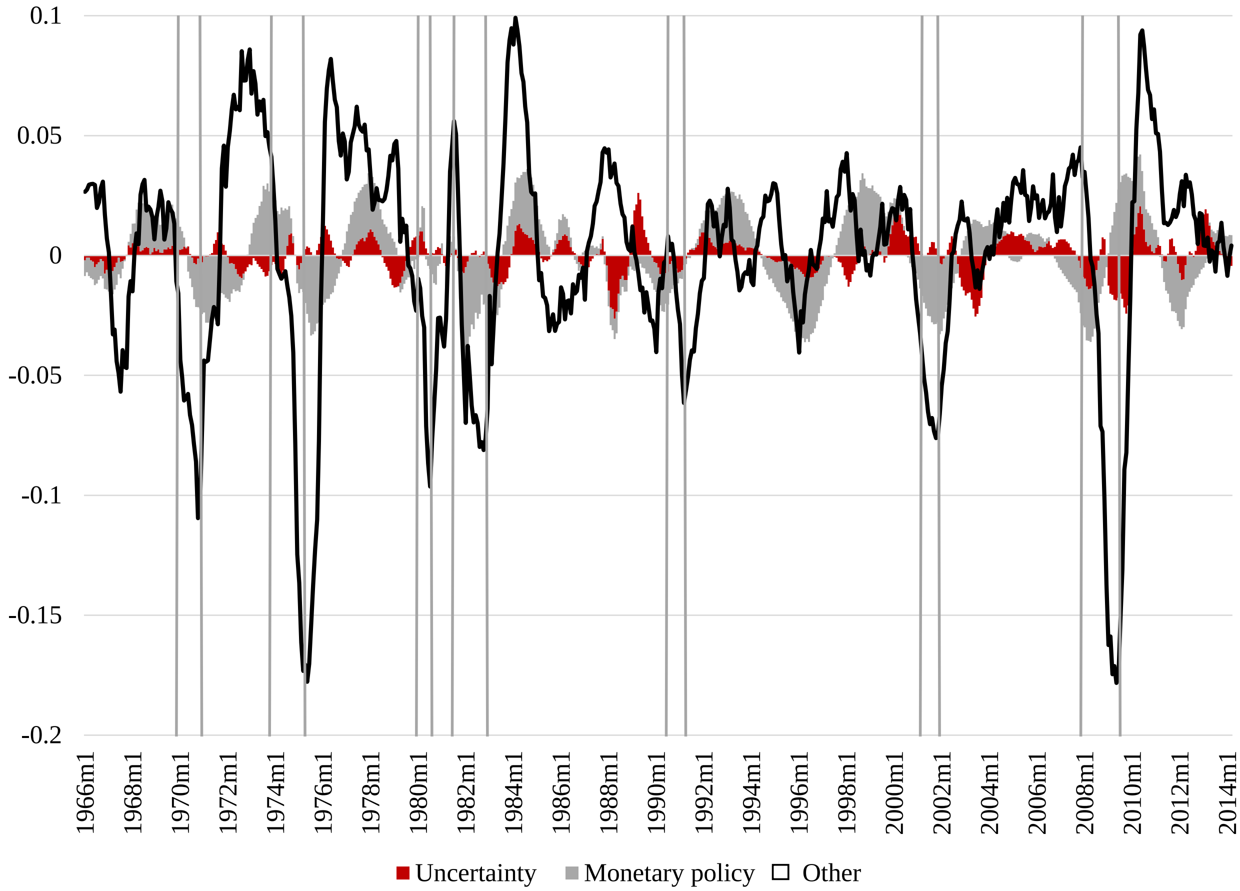
Notes: Dispersion measure is the IQR of sales growth of Compustat firms with 25+ years of data. The plotted cyclical components were obtained by the Baxter-King (1999) filter. The reported correlation and regression coefficients apply to the unfiltered time series.

Figure 1.4: Stock return dispersion over the business cycle



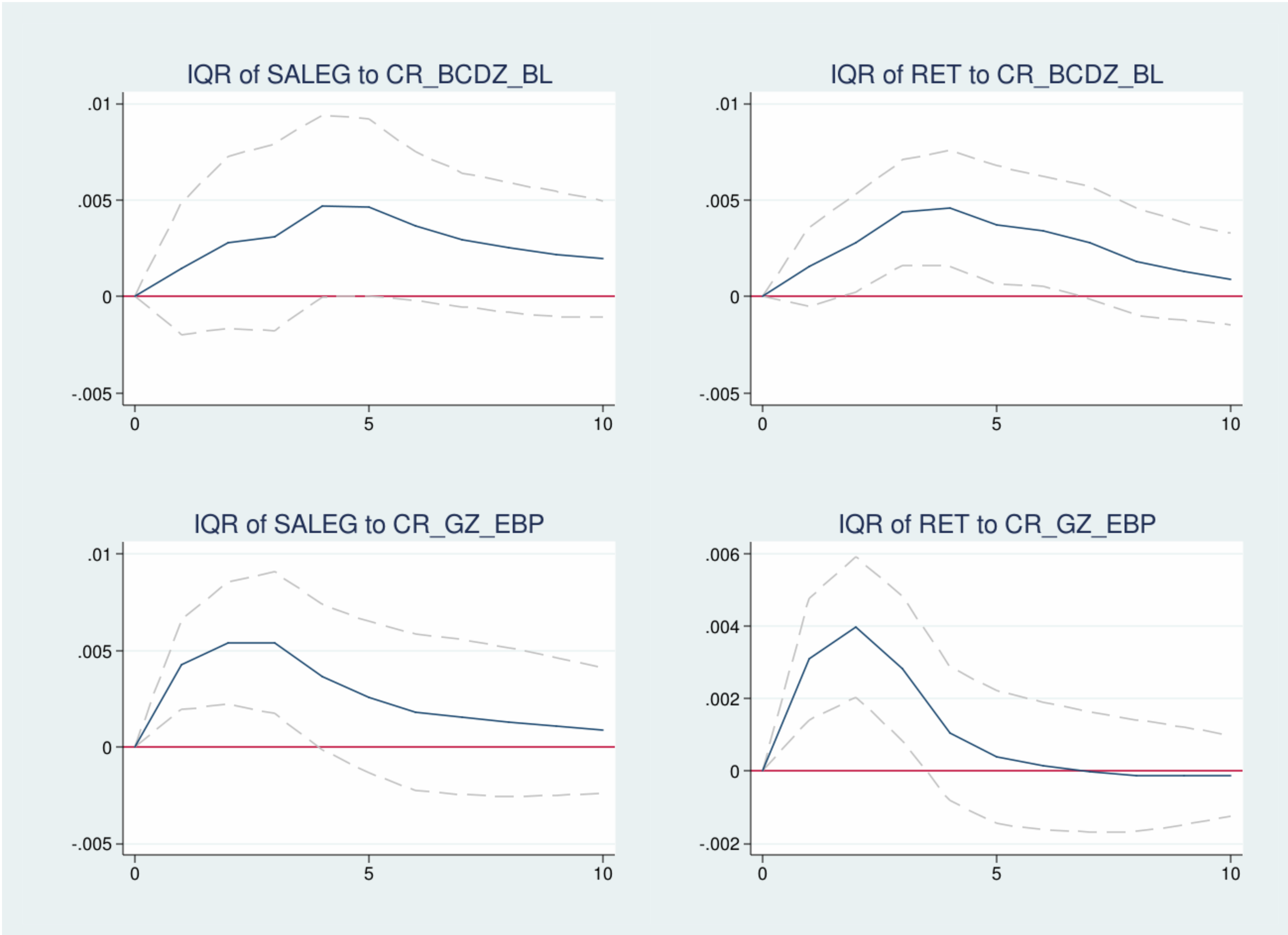
Notes: Dispersion measure is the IQR of stock returns of CRSP firms with 25+ years of data. The plotted cyclical components were obtained by the Baxter-King (1999) filter. The reported correlation and regression coefficients apply to the unfiltered time series.

Figure 1.5: Historical decomposition of industrial production in Bloom (2009) SVAR



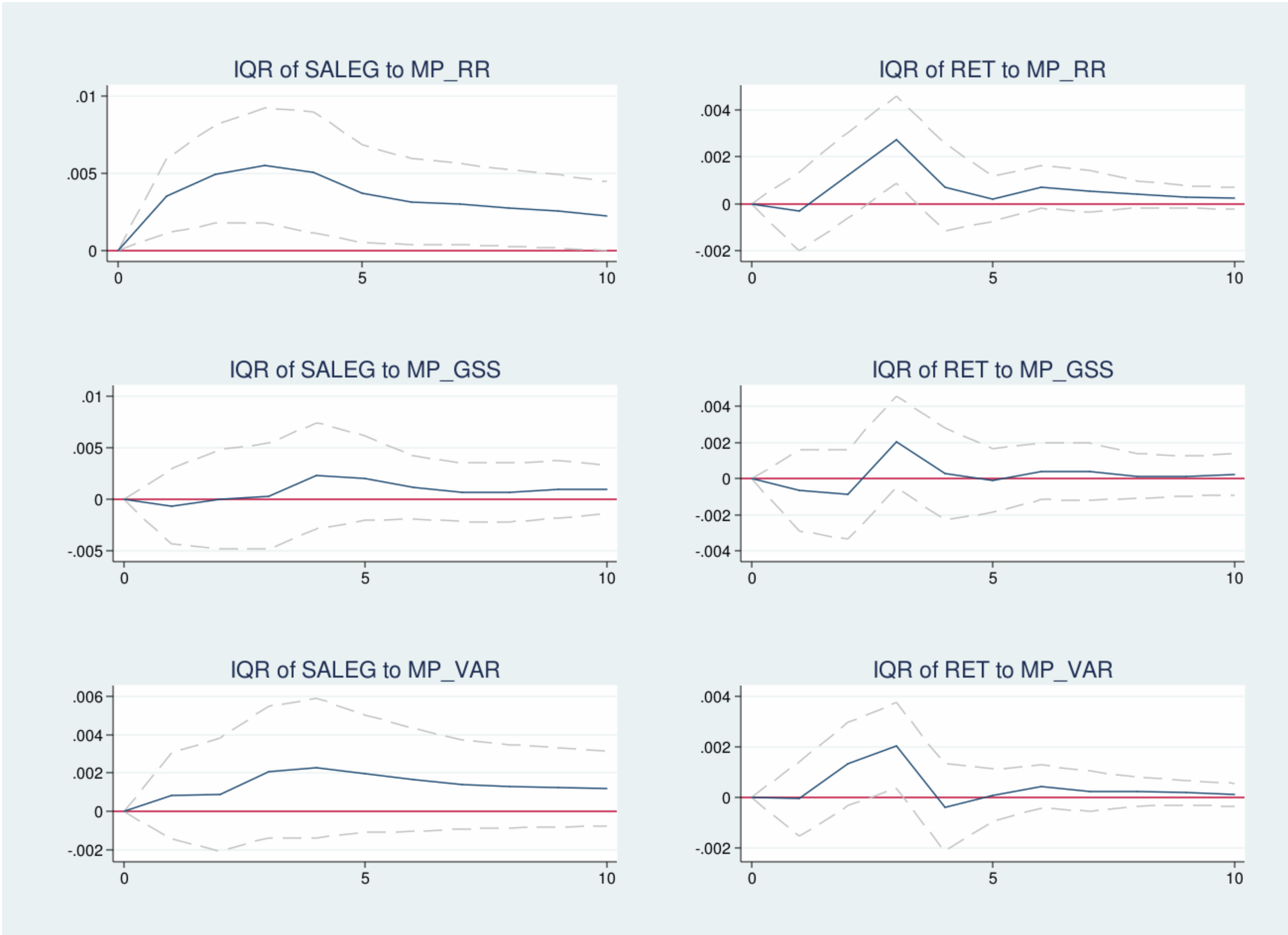
Notes: VAR estimated on monthly data from July 1962 through March 2014 with 12 lags. Variables: S&P500 stock-market index, VIX index (realized volatility pre-1986), Federal Funds rate, average hourly earnings for production workers (manufacturing), CPI (all urban consumers, seasonally adjusted), average hours in manufacturing (SA), employment in manufacturing (SA), industrial production in manufacturing (SA). All variables, except FFR and VIX, enter in logs. All variables are HP detrended using a filter value of $\lambda = 129600$. The thick black line shows the evolution of IP. The colored areas show the contribution of uncertainty shocks (innovations in VIX, red), monetary policy shocks (innovations in FFR, grey) and all other shocks (white). Structural innovations are identified through a Choleski decomposition with variable order given above. Vertical grey lines mark the beginning and end of NBER recessions.

Figure 1.6: Response of sales growth dispersion (left) and stock return dispersion (right) to **Credit Supply** shocks



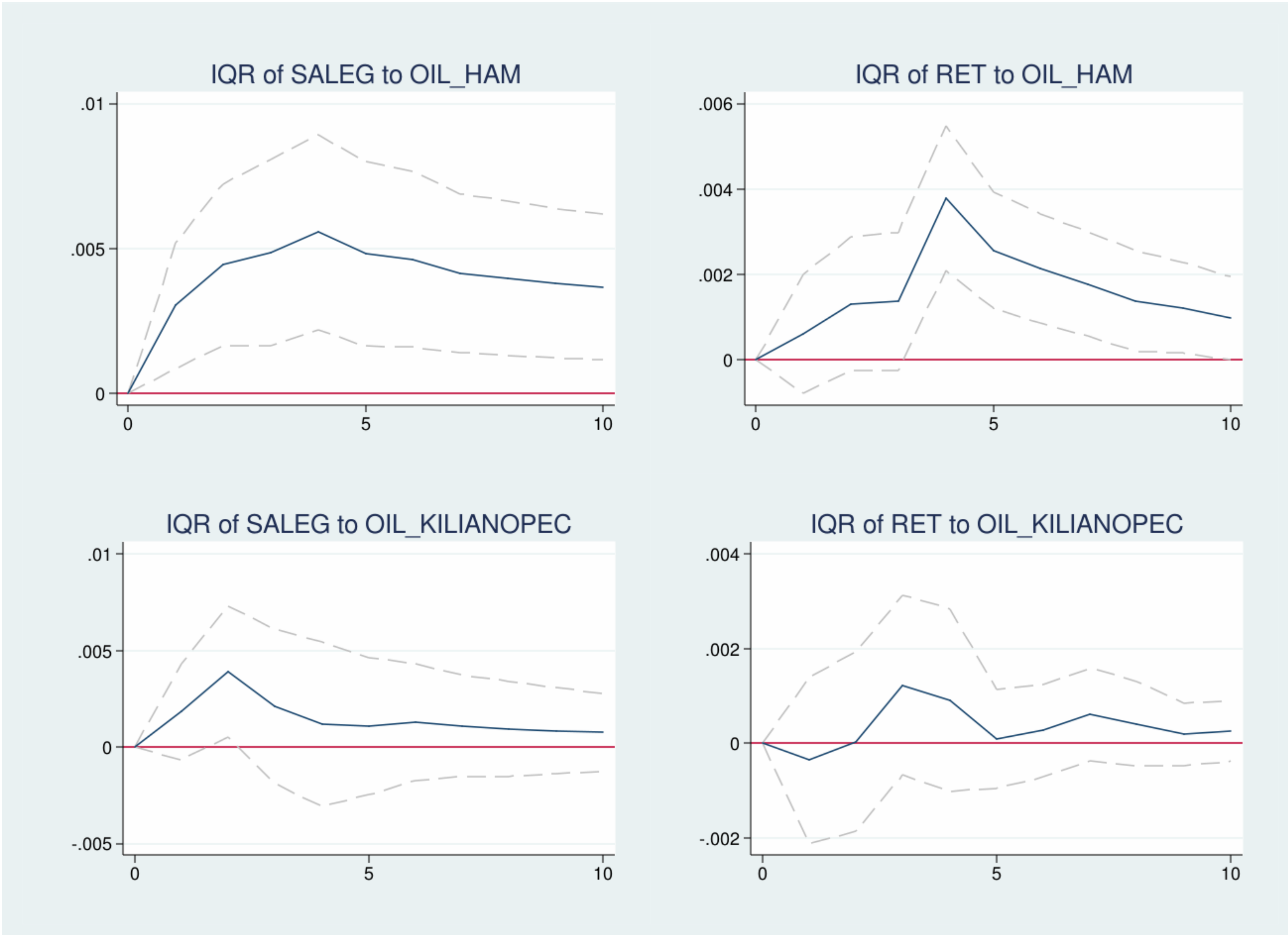
Notes: Implied impulse response functions from estimating equation (1.1) in the text. SALEG-sales growth; RET-stock returns. See Table 1.1 for the explanation of the credit supply shocks.

Figure 1.7: Response of sales growth dispersion (left) and stock return dispersion (right) to **Monetary Policy** shocks



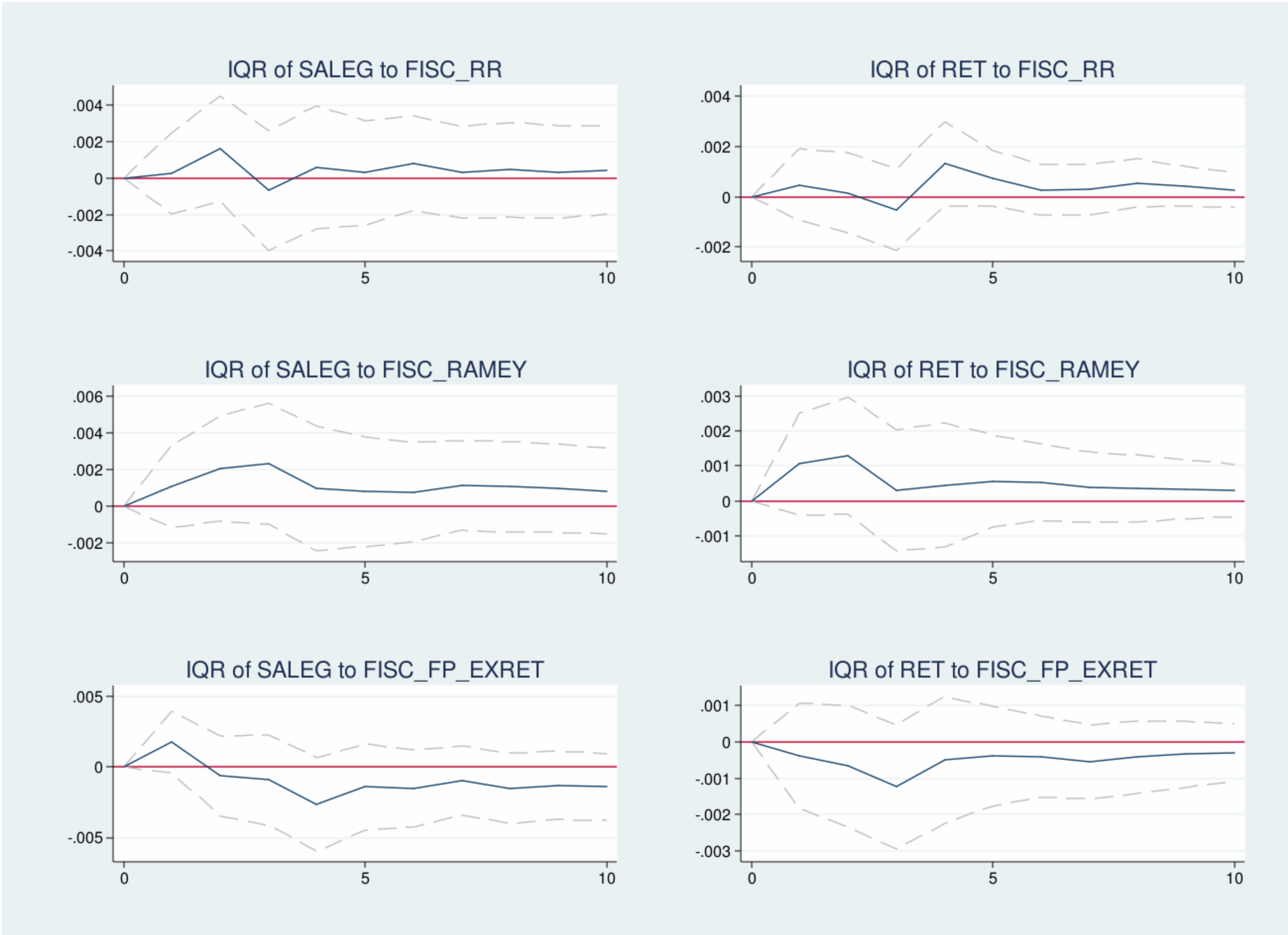
Notes: Implied impulse response functions from estimating equation (1.1) in the text. SALEG-sales growth; RET-stock returns. See Table 1.1 for the explanation of the monetary policy shocks.

Figure 1.8: Response of sales growth dispersion (left) and stock return dispersion (right) to **Oil Price** shocks



Notes: Implied impulse response functions from estimating equation (1.1) in the text. SALEG-sales growth; RET-stock returns. See Table 1.1 for the explanation of the oil price shocks.

Figure 1.9: Response of sales growth dispersion (left) and stock return dispersion (right) to **Fiscal** shocks



Notes: Implied impulse response functions from estimating equation (1.1) in the text. SALEG-sales growth; RET-stock returns. See Table 1.1 for the explanation of the fiscal policy shocks.

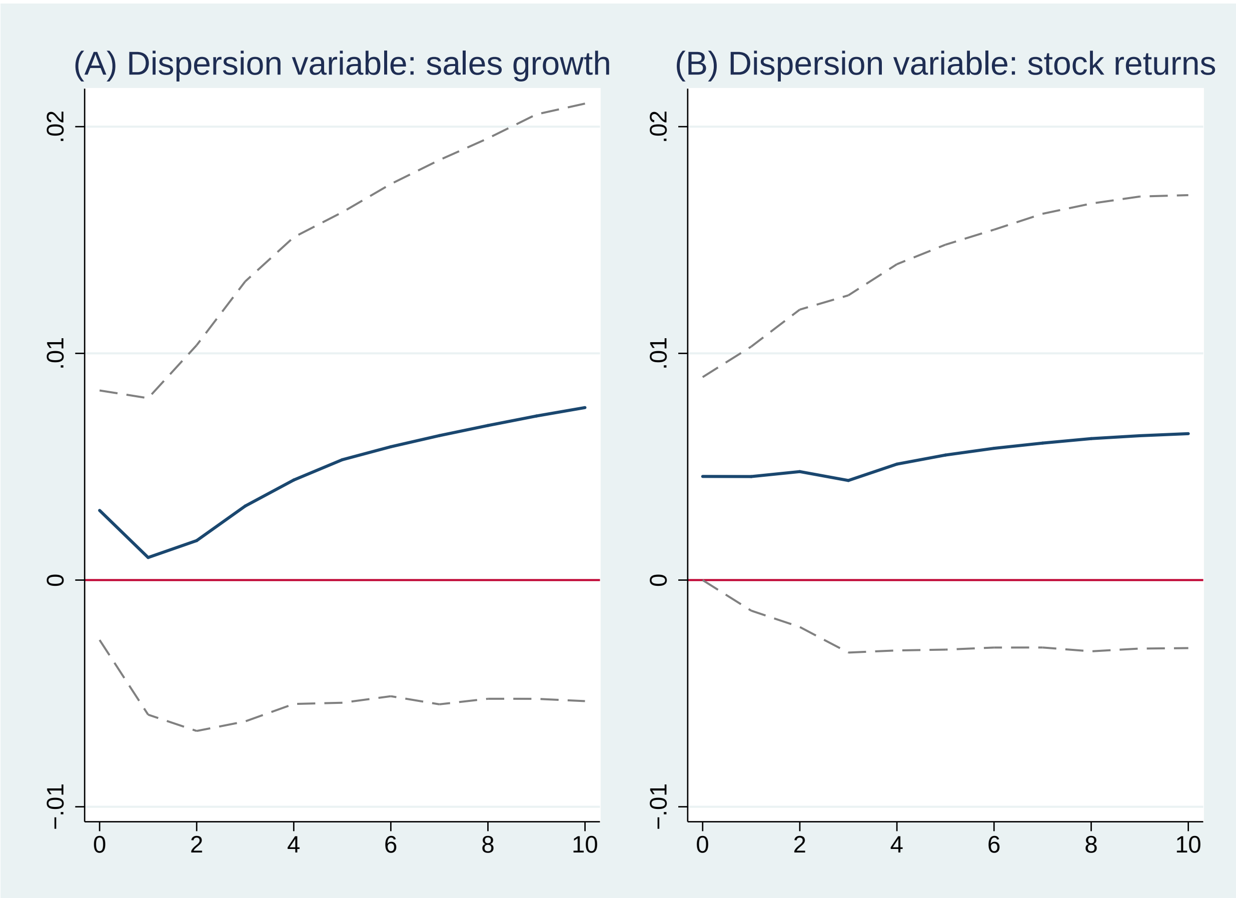
Figure 1.10: Response of sales growth dispersion (left) and stock return dispersion (right) to **Productivity** shocks

Notes: Implied impulse response functions from estimating equation (1.1) in the text. SALEG-sales growth; RET-stock returns. See Table 1.1 for the explanation of the aggregate productivity shocks.

Figure 1.11: Response of dispersion to an output level-shock (external instruments)

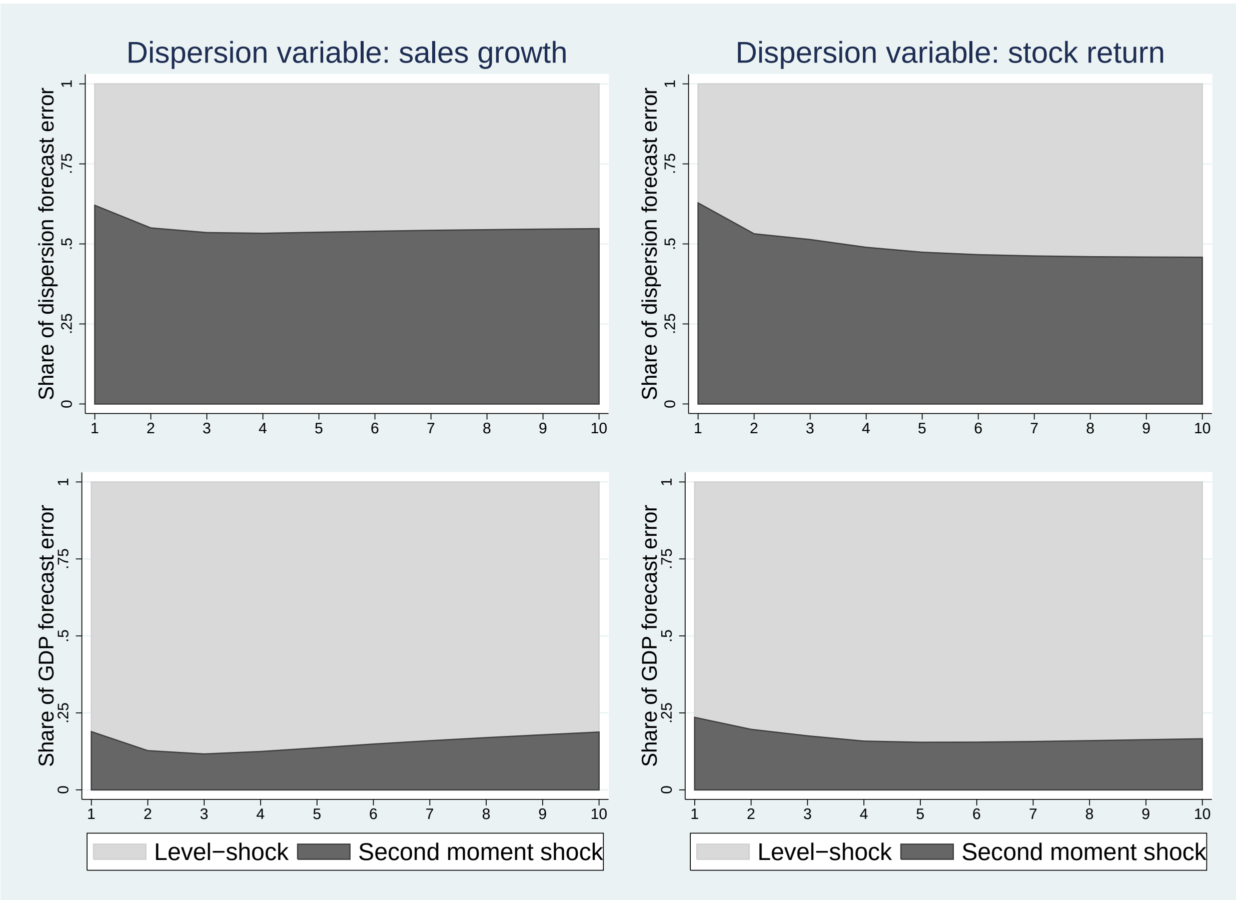
Notes: Two separate bivariate VARs are estimated on quarterly data from 1968Q1 through 2012Q4 with 4 lags. Variables: log GDP and sales growth dispersion (upper part); log GDP and stock return dispersion (lower part). Structural identification of the GDP level-shock is achieved by the Stock and Watson (2008) "external instruments" method (see Appendix A.2). The baseline estimation in Panel (A) utilizes all 5 instruments discussed in the text: Hamilton (1996) oil price shock; Romer and Romer (2004) monetary policy shock; Basu, Fernald and Kimball (2006) productivity shock; Fisher and Peters (2010) fiscal shock; Gilchrist and Zakrajsek (2012) credit supply shock. Panel (B) examines robustness by dropping the Gilchrist and Zakrajsek (2012) credit supply shock. Dashed lines indicate 95% Monte Carlo confidence bands. The output shock is normalized to -2 percentage points, which roughly corresponds to the fall in GDP in 2008Q4.

Figure 1.12: Response of output to a second-moment shock



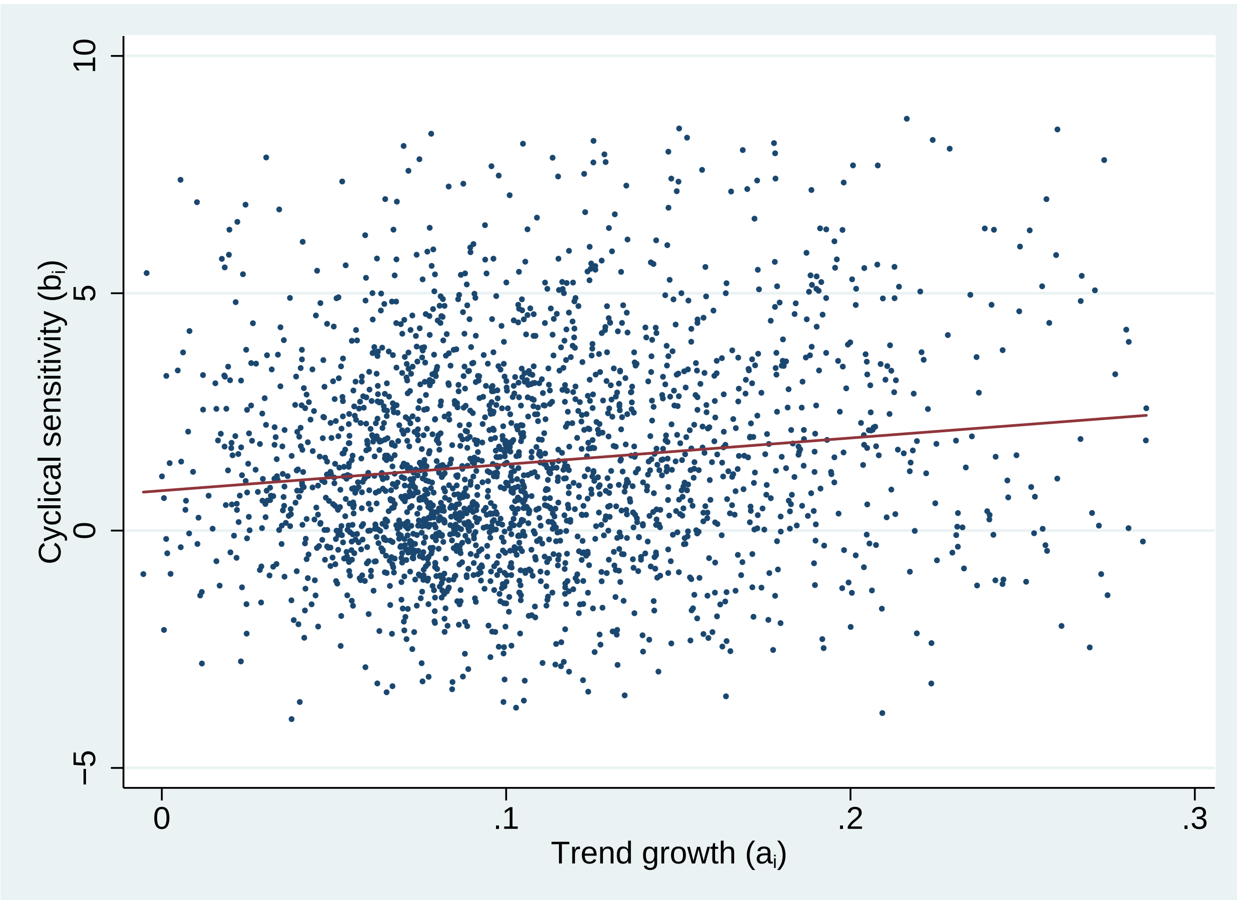
Notes: Two separate bivariate VARs are estimated on quarterly data from 1968Q1 through 2012Q4 with 4 lags. Variables: Panel (A)–log GDP and sales growth dispersion; Panel (B)–log GDP and stock return dispersion. The graphs show the IRF of output to an identified innovation in dispersion (a second-moment shock). Dashed lines indicate 95% Monte Carlo confidence bands. The shock is normalized to increase dispersion by 1 standard deviation. Structural identification of the output level-shock and the second-moment dispersion shock is achieved by the Stock and Watson (2008) "external instruments" method (see Appendix A.2). The estimation utilizes all 5 instruments of the baseline specification: Hamilton (1996) oil price shock; Romer and Romer (2004) monetary policy shock; Basu, Fernald and Kimball (2006) productivity shock; Fisher and Peters (2010) fiscal shock; Gilchrist and Zakrajsek (2012) credit supply shock.

Figure 1.13: Variance decomposition from the baseline SVARs



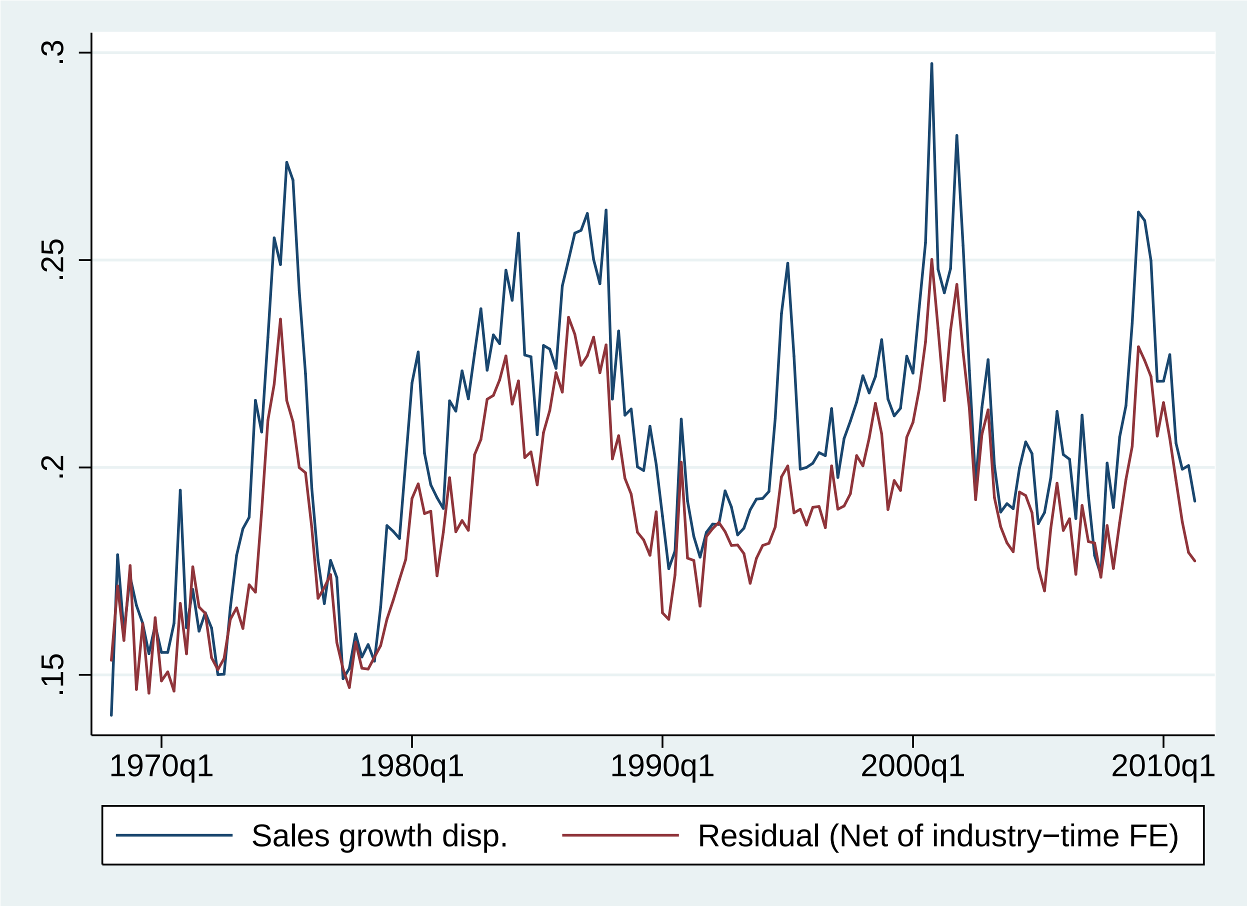
Notes: Two separate bivariate VARs are estimated on quarterly data from 1968Q1 through 2012Q4 with 4 lags. Variables: left panel–log GDP and sales growth dispersion; right panel–log GDP and stock return dispersion. The upper part shows the forecast error decomposition of the dispersion measures, and the lower part shows the decomposition of the output forecast error variance. Structural identification of the output level-shock and the second-moment dispersion shock is achieved by the Stock and Watson (2008) "external instruments" method (see Appendix A.2). The estimation utilizes all 5 instruments of the baseline specification: Hamilton (1996) oil price shock; Romer and Romer (2004) monetary policy shock; Basu, Fernald and Kimball (2006) productivity shock; Fisher and Peters (2010) fiscal shock; Gilchrist and Zakrajsek (2012) credit supply shock.

Figure 1.14: Testing the Abraham-Katz mechanism (sales growth)



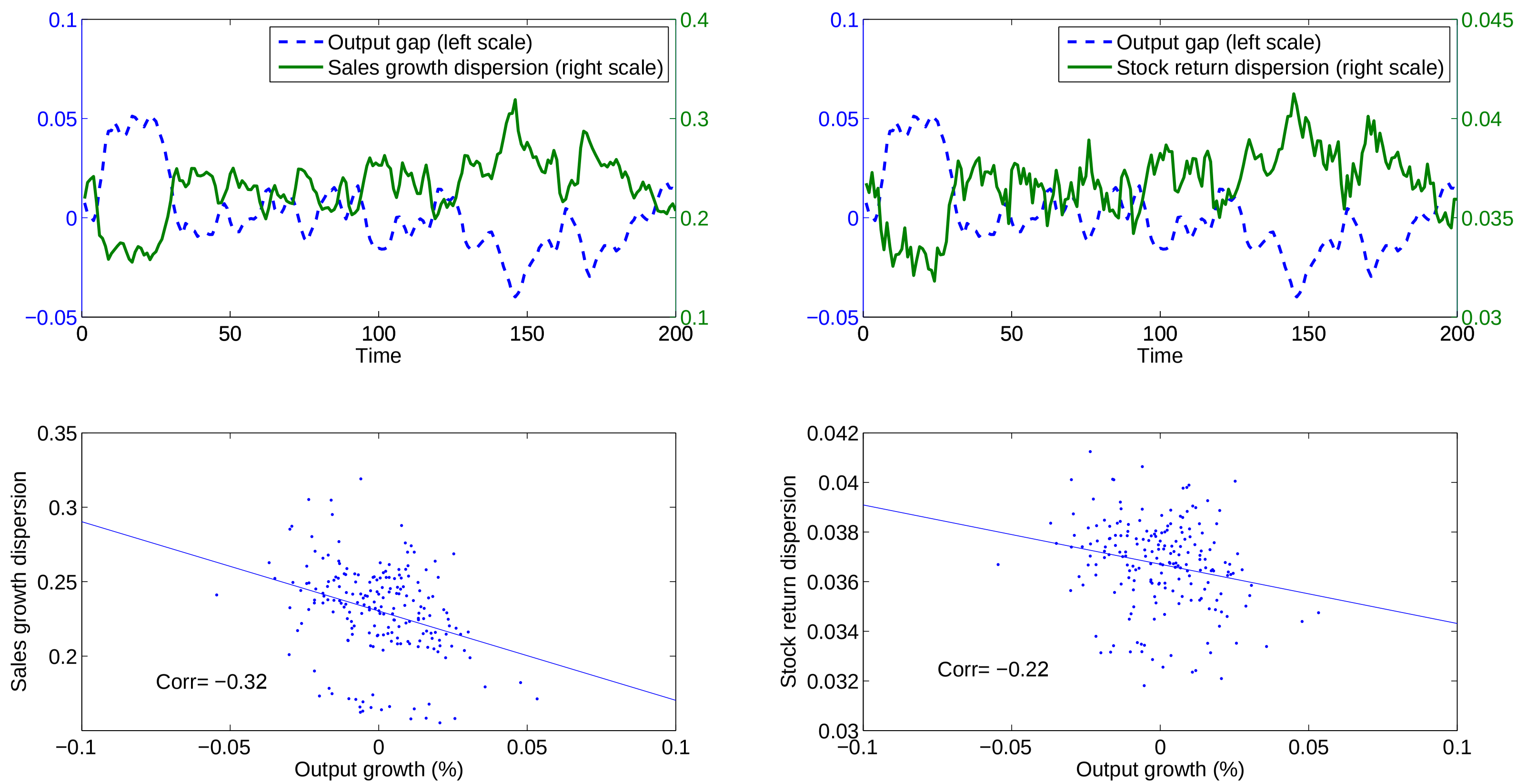
Notes: Results from estimating equation (1.2) in the text. Firm-level sales growth is regressed on detrended GDP growth and a constant. The constant is interpreted as the "trend growth" (a_i) and the coefficient on GDP growth is interpreted as the "cyclical sensitivity" (b_i) of each firm. As demonstrated by Abraham and Katz (1986), a negative correlation between trend growth and sensitivity could generate countercyclical cross-firm dispersion. However, the figure shows that this correlation is positive in the data.

Figure 1.15: Testing the sectoral shock hypothesis (sales growth)



Notes: Firm-level sales growth is regressed on a full set of industry and quarter fixed effects, and the cross-sectional dispersion of the residual is calculated. I use 41 industries: manufacturing at 3-digit NAICS codes, all other sectors at 2-digit NAICS codes. The residual dispersion series follows very closely the raw series (correlation 0.94). Sectoral shocks are unlikely the main source of countercyclical dispersion.

Figure 1.16: Model simulation: Dispersion and the business cycle



Notes: The model in Section 1.4 is simulated with 10,000 firms for 200 quarters. Preference parameters are calibrated to standard values (see section 1.5). Aggregate TFP is chosen to match the persistence and volatility of GDP. The underlying heterogeneity (α^{-1}) and the information cost function ($G(\beta)$) is calibrated to match *average* sales growth dispersion. The business cycle properties of sales growth dispersion (left) and stock return dispersion (right) emerge endogenously from the model. The volatility and correlation patterns of simulated dispersion time series are close to their empirical counterparts (see Figure 1.3 and 1.4).

Chapter 2

Inefficient Expectations: Assessing Imperfect Information Models Using Individual Forecasts

2.1 Introduction

Expectations are central to macroeconomic dynamics. How much to consume or save, what price to set, and when to invest are just some examples of forward-looking decisions that hinge on agents' expectations of the future. Yet our understanding of how those expectations are formed and how best to model them is very limited. Starting in the 1970s, researchers abandoned the mechanistic assumption of adaptive expectations in favor of model-consistent rational expectations. Rational expectations not only provided theoretical consistency, but also yielded many novel policy insights, which quickly rendered them the de facto standard in macroeconomics. Although early work in rational expectations modeling often emphasized search behavior and signal extraction problems in information acquisition (Phelps, 1968; Lucas, 1972), much of the subsequent theory has been built upon the assumption of full-information rational expectations (FIRE) on the part of all agents.

More recently, however, several empirical failures of full information models spurred renewed interest in the constraints that agents face when forming expectations. Sims (1998) points out that the joint dynamic behavior of prices, wages, and macroeconomic aggregates shows sluggish cross-variable responses that do not fit easily into any full information macroeconomic theories. "Classical" models imply that most kinds of disturbance to the economy should result in instantaneous movements in prices, while sticky price "Keynesian" models imply that the same shocks should produce instantaneous movements in output. Sims (1998) suggests that limited information-processing capacity can provide a common source of inertia in both prices and real variables. Building on this insight, recent work on rational expectations models with information frictions has shown that information rigidities can account for otherwise puzzling empirical findings (see Woodford, 2001; Sims, 2003; Mankiw and Reis, 2002; Mackowiak and Wiederholt, 2009). These theoretical models dispense with the assumption of full information, but they maintain rational expectations: although imperfect, agents use their available information efficiently, that is they form correct conditional expectations.

Recent research has used survey expectations to provide direct empirical evidence for the expectation formation process embedded in imperfect information macro models. Crowe (2010) and Coibion and Gorodnichenko (2012, 2015) recognize that matching the dynamic properties of aggregate time series is not a valid test of the microfoundations of models with information rigidities. Instead, they show that a large set of imperfect information models make clear predictions about the behavior of average forecasts across agents, and they use survey data on agents' expectations to test these predictions. In particular, they demonstrate that in models with information rigidities the average forecast error across agents responds slowly to shocks and it is predictable using past average forecast revisions. Since the survey data is consistent with these predictions about the average forecast, Coibion and Gorodnichenko (2015) concludes that "commonly observed rejections of the null of full-information rational expectations most likely reflect deviations from full information rather than departures from rational expectations."

The first contribution of this paper is to show that this conclusion based on the properties of consensus forecasts is misguided. The predictability of the average forecast error across agents is an emergent property in imperfect information models; that is, a property that arises only from the aggregation process and not at the individual level. At the individual level, the assumption of rational expectations demands that forecast errors are unpredictable. In turn, the consensus forecast may behave in the predicted way even if the actual individual forecasts are not obeying the restrictions imposed by rational expectations. Consequently, if we want to test the microfoundations of the expectations formation process in imperfect information macro models, it is preferable to use the microdata of individual forecasts.

Using a large panel of expectations data from the U.S. Survey of Professional Forecasters (SPF), I document that individual forecasts robustly violate rational expectations. Professional forecasters are some of the most informed economic agents, and the SPF is widely used in gauging expectations about the future path of the economy. Yet, a consistent pattern across different target variables emerges: a survey respondent's forecast revision, the difference between his current and previous forecast, has statistically and economically significant predictive power for both his subsequent forecast revisions and his forecast error itself. Since a forecaster's previous forecast is clearly in his information set, the presence of information rigidities at the individual level cast doubt on the microfoundations of rational expectations models with information frictions. In the language of noisy information models, there is pervasive evidence that something goes wrong already at the individual signal extraction step.

To account for the documented behavior of both the consensus forecast and individual expectations, I propose a simple modification of the noisy private information framework employed in the literature. This modification allows a certain form of non-rationality in the expectations formation process. Instead of calculating mathematically correct conditional expectations, agents use heuristic linear forecasting rules when they update their expectations after receiving new signals. I do not restrict the coefficients of the linear forecasting rules to be optimal in the minimum MSE sense, but I treat them as primitive parameters that need to be estimated from the data. Relaxing only this assumption, my framework nests the setup of earlier empirical work and it aligns with the information environment of many theoretical models. Furthermore, I show that three-dimensional panel data on the time path of individual agents' forecasts for different horizons makes it possible to estimate the parameters of this heuristic signal extraction model.

I estimate the model parameters using forecasts from the SPF for 13 different target variables. I find that in most cases agents seem to put too much weight on their private information compared to the MSE minimizing weight, which may be interpreted as overconfidence in the precision of private information. Looking across target variables, there is no systematic relationship between the optimal and the heuristic weight, which further suggests that rational expectations is not a useful approximation of the expectations formation process. The richness of the SPF data allows me to perform a battery of consistency checks on the assumed model of information processing. I derive additional time-series (within forecaster) and cross-sectional (between forecaster) implications of the basic setup, and I show that the estimates from these alternative specifications line up reasonably well with my baseline results. This confirms that my model of the forecasting process, although still very stylized, is able to match the survey data along some important dimensions. Thus, the model can be a useful tool to examine the deviations from the rational benchmark in the expectations formation process.

Finally, I study the cross-sectional heterogeneity in the degree of forecaster irrationality. First, I uncover substantial differences in how much forecasters deviate from the rational benchmark. Second, I find that having a relatively low or relatively high individual information rigidity is largely an inherent property of the forecaster. That is, a forecaster who exhibits over-reaction to new information when forecasting GDP growth also tends to exhibit over-reaction when forecasting inflation, and vice versa. Exploiting that the SPF collects forecasts for several target variables from each respondent, I use principal component analysis to construct two composite measures of individual information rigidity. The two independently derived measures are highly correlated across forecasters, which provides a further robustness check on the statistical model of the expectations formation process. Third, I try to relate individual information rigidity to some observable characteristics of the forecaster, such as the level of experience and a measure of information quality.

The last part of the paper discusses potential interpretations of my empirical results and their implications for macroeconomic theory. On the one hand, it is possible that survey respondents have non-standard objective functions, and thus they do not report their true expectations in the SPF. Using a simple game as in Morris and Shin (2002), I show that reputational concerns may lead agents to under- or overweight private information compared to the mathematical expectation. On the other hand, if we interpret SPF forecasts as agents' genuine expectations, then my results provide evidence against the microfoundations of models where agents form rational expectations from noisy signals. It is important to emphasize, however, that the failure of agents to use the mathematically correct forecasting rules do not undermine the most important aggregate implications of imperfect information macro models. The main mechanism in these models is the gradual discovery of the true state of economy from dispersed information, which can be imposed directly as a behavioral assumption.

This paper is related to the long literature on the rationality of both individual and consensus forecasts in various expectations surveys. Pesaran and Weale (2006) provide an extensive survey. This paper differs in two important ways from most the work in this literature. First, instead of focusing on a specific target variable, I use forecasts for many variables and horizons both at the aggregate and at the individual level. Using a rich dataset allows me to draw conclusions that are more general. Second, I rely on a model of the expectations formation process to guide my choice of the relevant regressors when testing the rational expectations hypothesis. Interpreting the regression coefficients through the lens of this model also provides a precise metric to quantify the deviations from rational expectations.

This paper is most directly related to the recent empirical work that exploits survey data on agents' forecasts to assess the relevance of imperfect information rational expectations models. Mankiw, Reis and Wolfers (2004) assess whether a sticky-information model can replicate some stylized facts about the predictability of inflation forecast errors. Crowe (2010) analyzes consensus forecasts for GDP growth across many countries to test the implications of noisy information models. Two closely related papers by Coibion and Gorodnichenko (2012, 2015) study the evidence for imperfect information models using consensus forecasts and disagreement measures from U.S. and international surveys. My paper contributes this literature in a number of ways. First, instead of using different moments of the survey forecasts, such as the average or standard deviation, I use the forecaster-level microdata directly to assess the models' microfoundations. Second, I examine whether different regression specifications derived from the model yield similar estimates for the underlying parameters. Third, the microdata allows me to study the heterogeneity in forecaster rationality.

The rest of the paper is structured as follows. Section 2.2 presents the tests of forecast rationality, and introduces a simple model to explain and measure the deviations from rational expectations. Section 2.3 derives further predictions from the model, and confronts them with the survey data. The cross-sectional heterogeneity in forecasters' informational efficiency is analyzed in Section 2.4. Section 2.5 considers different interpretations of the reduced form results, and closes with a discussion of the implications for macroeconomic theory. Some technical derivations are relegated to the Appendix.

2.2 Forecast errors, forecast revisions and models of expectations formation

Using survey expectations, this section provides pervasive evidence of a systematic relationship between forecast errors and forecast revisions both for the consensus (average) forecast and for individual forecasters. Earlier empirical work presented the sluggish adjustment of the consensus forecast as evidence for rational models with information frictions. However, the individual results are at odds with this interpretation. This section introduces a simple model of the expectations formation process that can jointly account for the aggregate and the micro-level stylized facts. I estimate the model, and investigate the deviations from the rational expectations benchmark.

Data

I use data from the Survey of Professional Forecasters for several reasons. Since professional forecasters are some of the most informed agents in the economy, they provide a conservative benchmark for assessing possible deviations from rational expectations. Second, the microdata of professional forecasters' predictions are consistently available at a quarterly frequency for several target variables and time horizons. This amount of detail is unique among macroeconomic expectations surveys, and it is essential for most of the analysis in the paper. Third, professional forecasters, unlike consumers in other surveys, make predictions of explicitly defined macroeconomic variables, such as the CPI or GDP; thus, there is a well-defined relationship between their forecasts and ex-post values. Finally, aggregate data on professional forecasters have been used in support of imperfect information theories.

The SPF is a quarterly survey of approximately 30-40 professional forecasters. Forecasts are collected by the Philadelphia Federal Reserve in the middle of each quarter for a variety of macroeconomic variables. The non-overlapping forecasting horizons range from the current quarter to four quarters ahead. Forecasts for real GDP, housing starts, industrial production, the GDP price deflator and the unemployment rate are available from 1968Q4. These variables have the longest time series and the highest coverage by forecasters, so I will refer to them as "core variables" in the analysis. Starting in 1981Q3, the SPF also includes forecasts of 8 additional macroeconomic variables: the 3-month Treasury bill rate, Moody's AAA corporate bond yield, the overall CPI, real consumption expenditures, real federal government expenditures, real state and local government expenditures, real nonresidential investment, and real residential investment. For the NIPA series, industrial production and the CPI, I construct forecasts of annualized quarterly percent changes from the underlying forecasts of the levels, while housing starts, unemployment, and the interest rate variables enter the analysis in their reported units. Forecast errors are constructed using real-time data that was available one year after the forecast target period. I use historical vintages from the Philadelphia Fed's Real-Time Data Set to measure ex-post variables because final data may reflect reclassifications and redefinitions such that the final values are not directly comparable to the historical forecasts made by agents (Croushore and Stark, 2001). Table 2.1 provides some descriptive statistics about the size of the dataset used in this paper.

Tests of FIRE at the aggregate and individual levels

Suppose that survey respondents want to forecast the future values of a macroeconomic variable z_t. The h-period ahead forecast of agent i made at time t is denoted by $z_{i,t+h|t}$. Let us decompose the individual information sets, Ω_{it}, into a public information set Ψ_t, and an individual-specific private information set Φ_{it} such that $\Omega_{it} = \Psi_t \cup \Phi_{it}$. A strict interpretation of full-information rational expectations has the obviously unrealistic implication that agents have identical expectations about the future: $z_{i,t+h|t} = E\left[z_{t+h}|\Psi_t\right]$ for all i. Although forecaster disagreement is prevalent in the data, it is still standard practice to test whether consensus forecasts are consistent with FIRE. The idea is that even if individual forecasters do not satisfy FIRE, the average forecast can still be a good approximation for the expectations of the fully informed representative agent in theoretical models. Commonly employed tests in the literature make use of the "orthogonality condition" of expectations errors under FIRE: $E[z_{t+h} - \overline{z_{t+h|t}}|I_t] = 0$ where $I_t \subseteq \Psi_t$ and $\overline{z_{t+h|t}}$ is the average forecast across agents. In practice, researchers regress the consensus forecast error on a subset of the information available at the time the forecast was made, and check if the error is predictable.

Table 2.2 presents results from test regressions in which the conditioning variable is the ex-ante forecast revision:

$$z_{t+h} - \overline{z_{t+h|t}} = c + \beta\left(\overline{z_{t+h|t}} - \overline{z_{t+h|t-1}}\right) + \varepsilon_{th} \tag{2.1}$$

for $h = 0, .., 4$. In most cases, the informational efficiency of consensus forecasts is strongly rejected in the SPF. For 10 out of the 13 macroeconomic variables, including all of the core variables, the coefficient on the forecast revision is positive and statistically significant. There is only one negative point estimate but it is highly insignificant.[1] The regressions suggest that the consensus forecast adjusts only partially to new information; as a result, the forecast revisions have predictive power for the forecast errors. As the next section demonstrates, the sluggish adjustment of the consensus forecast is consistent with imperfect information models, which have been framed in recent empirical work as supporting evidence for these models. However, the model assumptions are formulated in terms of the expectations formation of individual agents and not in terms of the average expectations across agents. This suggests that we should study individual forecasts when evaluating the empirical relevance of the theories.

An analogous test on individual survey responses reveals that forecasters in the SPF do not have rational expectations. Table 2.3 reports pooled regressions of individual forecast

[1] Since the regression errors potentially have a complicated correlation structure due to the three dimensions of the data, I construct all standard errors in the paper by the method of Driscoll and Kraay (1998). This estimator does not require strong assumptions on the form of the cross-sectional and temporal correlation in the error terms. The structural model of the expectation formation process presented in section 2.2 will confirm this precaution.

errors on individual forecast revisions. For 12 out of the 13 target variables, the forecast revision is a highly significant predictor of the forecast error. In contrast to the aggregate results, the coefficient is negative in most cases, so an ex-ante upward revision signals that the forecast will be too high ex-post, and vice versa. These results provide a very powerful rejection of rational expectations, because they require minimal assumptions about agents' information sets. Standard regression-based tests of rationality are plagued with the issue of what information was available to the forecaster at the time the forecast was made and how well the forecaster understood that information. It is straightforward to assume that the agent knows his previous forecast when he forms expectations at time t.

Noisy information with rational and heuristic filtering

Coibion and Gorodnichenko (2015) argue that the behavior of the consensus forecast is compatible with rational expectations in the presence of information frictions. I lay out a simple setup to illustrate their point. Assume that a unit mass of rational agents wants to forecast a macroeconomic variable. The variable follows an AR(1) process:

$$z_t = \rho z_{t-1} + v_t, \quad 0 \leq \rho \leq 1$$

where $v_t \sim iidN(0, \sigma_v^2)$. Agents face an information friction: they cannot directly observe z_t, but instead receive a private signal y_{it} such that

$$y_{it} = z_t + \omega_{it}, \tag{2.2}$$

where $\omega_{it} \sim iidN(0, \sigma_\omega^2)$ represents the noise in the signal. Each agent i forms Bayesian forecasts using the Kalman filter. If $z_{i,t+h|t} = E\left[z_{t+h} \mid y_{it}, y_{i,t-1}, y_{i,t-2}, ...\right]$, then we can write

$$z_{it|t} = G^* y_{it} + (1 - G^*) z_{it|t-1}, \tag{2.3}$$

$$z_{i,t+h|t} = \rho^h z_{it|t}, \tag{2.4}$$

where G^* is the Kalman gain which represents the weight placed on new information relative to previous forecasts. Equation (2.3) describes how agents update their beliefs about the current value of z_t, while equation (2.4) uses the AR(1) structure to iterate forward this belief. The value of G^* is determined by the persistence of the series (ρ) and the noise-to-signal ratio ($\sigma_\omega^2/\sigma_v^2$).

To see the implications for the consensus forecast, let $\overline{z_{t+h|t}} = \int_0^1 z_{i,t+h|t} di$ be the average h-step ahead forecast among all agents. In Appendix B.1, I show that the following relationship between ex-post average forecast errors and ex-ante average forecast revisions holds:

$$z_{t+h} - \overline{z_{t+h|t}} = \frac{1 - G^*}{G^*} \left(\overline{z_{t+h|t}} - \overline{z_{t+h|t-1}} \right) + v_{t+h,t}, \tag{2.5}$$

where $v_{t+h,t} = \sum_{j=1}^{h} \rho^{h-j} v_{t+j}$ is the rational expectations error. The error term is uncorrelated with time t information, so the β parameter in the OLS regression of (2.1) consistently estimates $(1 - G^*)/G^*$. The coefficient on average forecast revisions maps directly to the Kalman gain, and it can be interpreted as the relative importance of old information (the prior) as opposed to new information (the signal) when agents form expectations. If the signal is completely revealing ($\sigma_\omega^2 = 0$), then $G^* = 1$ and $\beta = 0$. When the signal is very noisy ($\sigma_\omega^2 \to \infty$), G^* is close to zero and $\beta \to \infty$. Since $v_{t+h,t}$ has a complex correlation structure across survey dates and forecast horizons, the Driscoll and Kraay (1998) standard errors in earlier tests are justified. In line with the regression results from the SPF, equation (2.1) implies that the average forecast error is predictable. However, this does not contradict rational expectations. Agents cannot observe the forecasts of others, so the average forecast and average revision are not part of their individual information sets. Intuitively, noisy information causes rational agents to discount their private signals. The partial reaction of individual forecasters implies that the consensus forecast also reacts only partially to the aggregate signal. However, the aggregate signal – being the average of individual signals – has much less noise than the signal of any single individual. Thus, the consensus forecast places an inefficiently low weight on the aggregate signal.

Unfortunately, the microdata underlying the consensus forecast reveals that rationality of expectations already fails at the individual level. Optimal signal extraction from noisy information cannot explain the individual regression results. To account for the predictability of forecast errors at both the aggregate and individual level, I propose a simple modification of the expectations formation process outlined above that allows a certain form of non-rationality. First note that the derivation of equation (2.5) did not rely on G^* being chosen optimally in a Bayesian sense. With any arbitrary G, the same relationship between average forecast errors and average forecast revisions would emerge as long as agents use an updating rule like the one in (2.3). We can interpret such a heuristic updating rule as a behavioral analogue to the Kalman filter. Due to parameter uncertainty or computational complexity, agents may use a simple linear rule when tackling the signal extraction problem.[2] With $G \neq G^*$, however, the orthogonality condition for individual forecast errors is not satisfied, and the model is consistent with the results in Table 2.3.

It can be shown that the following relationship between individual forecast errors and forecast revisions holds in the model:

$$z_{t+h} - z_{i,t+h|t} = \overbrace{\frac{1-G}{G}}^{\beta} \left(z_{i,t+h|t} - z_{i,t+h|t-1} \right) \overbrace{-\rho^h \omega_{it} + v_{t+h,t}}^{\varepsilon_{it}}. \tag{2.6}$$

This regression equation, of course, would be misspecified under the assumptions of my setup. The error term contains ω_{it}, which also factors into the forecast revision as demonstrated by (2.3). Appendix B.2 works out the asymptotics of the OLS estimator and shows that

$$\widehat{\beta}_{OLS} \xrightarrow{p} \beta + \frac{Cov(z_{i,t+h|t} - z_{i,t+h|t-1}, -\rho^h \omega_{it} + v_{t+h,t})}{Var(z_{i,t+h|t} - z_{i,t+h|t-1})} = f(G, \rho, \sigma_\omega^2/\sigma_v^2), \tag{2.7}$$

[2] This assumption is similar to the "natural expectations" assumption in some recent theoretical macro models. See, for example, Fuster, Hebert and Laibson (2012).

where $f(G, \rho, \sigma_w^2/\sigma_v^2)$ is strictly decreasing in G. A forecaster with rational expectations would choose a weight G^* that ensures $f(G^*, \rho, \sigma_w^2/\sigma_v^2) = 0$. Thus, if the regression coefficient on ex-ante forecast revisions is positive (negative), we must conclude that the forecasters put too little (too much) weight on new information.

Analyzing the expectations data and the documented failures of FIRE through the lens of this heuristic filtering model has several attractive features. First, it constitutes a minimal deviation from the framework of much existing work that assumes rational expectations. For example, many theoretical papers use an informational environment in which agents receive noisy private signals about an exogenously evolving macroeconomic variable such as nominal aggregate demand (Woodford, 2001; Mankiw and Reis, 2002; Mackowiak and Wiederholt, 2009). Relevant empirical work also starts from this setup (Coibion and Gorodnichenko, 2012, 2015). It is constructive to see how far we can go in matching the survey data by relaxing only one assumption. Second, because this setup nests the rational expectations benchmark, it can also be used to quantify the deviation from rational expectations. Third, as Section 2.3 demonstrates, the model imposes a number of testable restrictions on the data which allows for robustness and consistency checks.

Estimating informational inefficiencies

The parameters of the model can be estimated using the panel data on individual expectations in the SPF and the time series of the target variable. Under my assumptions, I can also recover the optimal weight that Bayesian agents would put on new information. Thus, we can define the deviation from fully rational expectations as $G - G^*$, that is the difference between the heuristic weight on new information and the optimal Kalman gain. Comparing this measure across target variables, we can gain additional insights into the applicability of rational expectations as an approximation of the expectations formation process. For each of the 13 macroeconomic variables in my dataset, I perform the following steps. First, I fit a first-order autoregressive process which yields an estimate of the persistence of the series. Second, I use the consensus regressions reported in Table 2.2 to uncover an estimate of G. Third, given my estimates of ρ and G, I use (2.7) to calculate the implied noise-to-signal ratio $(\sigma_\omega^2/\sigma_v^2)$ from the individual regressions reported in Table 2.3. Finally, I solve for the mathematically correct Kalman gain G^* and compare it to the heuristic weight G.

Can we learn anything new from this exercise? After all, the individual regressions already revealed that expectations are not informationally efficient, and one could simply use the regression coefficient on the forecast revision to gauge the "degree of irrationality." For example, Dovern et al. (2015) use this coefficient to compare the optimality of growth forecasts in the Consensus Economics survey across countries. My approach has two advantages. First, I provide a structural interpretation of the numbers. Any deviation from rational expectations comes from a suboptimal weighting of new information against agents' prior. Second, and more importantly, the coefficient on individual forecast revisions might not be a sufficient statistics to compare the inefficiency of forecasts. As demonstrated by (2.7), the OLS coefficient in the individual regression depends on several parameters. Without knowing these parameters, we cannot infer from the regression coefficient how far we are from rational expectations and direct comparison of the regression coefficients is meaningless. We need to use the structure of the model and bring in additional information to interpret the reduced form results.

Figure 2.1 compares the optimal and heuristic weights. There is substantial heterogeneity in both G and G^* across target variables. The rational expectations benchmark corresponds to the 45-degree line; forecasts for variables above the line tend to overreact to new information, while forecasts for variables below the line tend to underreact to news. Importantly, there seems to be no relationship between the optimal weight and the heuristic weight for this 13 macroeconomic variables. If the heuristic updating rules at least had a tendency to move together with their Bayesian counterparts, then rational expectations could still be a useful approximation of the expectations formation process. My results do not support such an interpretation.

2.3 Robustness and consistency checks on the model

This section derives a number of additional predictions of the model, and confronts them with the data. First, I test some theoretical restrictions on the coefficients of the forecast error-forecast revision regressions presented earlier. Second, I show that forecaster disagreement is directly related to the noise-to-signal ratio in the model, and I compare this cross-sectional moment with the time series estimates. Third, I derive an alternative regression specification to estimate agents' updating rules. This specification does not rely on the ex-post values of the forecasted variable, so it provides an independent way to identify the weight parameter. Although my model of expectations formation is clearly very stylized, these robustness checks point to a fairly high degree of consistency between model predictions and the data.

Coefficient restrictions in the baseline regressions

First, the model predicts a constant of zero in (2.1). As Table 2.2 shows, we cannot reject this restriction for the majority of target variables. Interestingly, the two exceptions are interest rate variables (T-bill and corporate bond rates) where there seems to be a systematic upward bias in the forecasts. The same restriction applies to the individual level regressions in Table 2.3. Again, in most cases we cannot reject that the constant is zero. Second, the model implies that the inclusion of forecaster fixed effects in the pooled individual regressions should not change the coefficient on the forecast revision. When I re-estimate the regressions with fixed effects, the estimated values of β and the standard errors remain essentially unchanged as expected (Table 2.4). Third, the model predicts that the coefficients on the contemporaneous forecast and on the lagged forecast in (2.1) are equal in absolute value. To implement this additional test, I decompose the consensus forecast revision into two terms as follows

$$z_{t+h} - \overline{z_{t+h|t}} = c + \beta_1 \overline{z_{t+h|t}} + \beta_2 \overline{z_{t+h|t-1}} + \varepsilon_{th}. \quad (2.8)$$

Under my model, we expect $\beta_1 > 0, \beta_2 < 0$, and $\beta_1 + \beta_2 = 0$.[3] Table 2.5 presents the results from estimating equation (2.8) for the 13 target variables. The signs of both coefficients conform to the theoretical predictions for almost all variables, and for the two exceptions the point estimates are insignificant. Moreover, for most target variables, including all of the core variables, we cannot reject the null that the sum of the two coefficients is equal to zero. The results thus provide additional evidence that the expectations formation process of professional forecasters can be approximated as a heuristic signal extraction problem.

Cross-sectional predictions

Forecasters in the model are identical except for their private information. Any disagreement in forecasts comes from the different signals that agents receive. This implies a close connection between the noisiness of individual forecasters' private information and the disagreement across forecasters. Appendix B.3 shows that the cross-sectional variance of forecasts submitted for the current quarter is given by[4]

$$Var_i\left(z_{it|t}\right) = \frac{G^2\sigma_\omega^2}{1-\left(1-G\right)^2\rho^2}. \tag{2.9}$$

As a consistency check, we can obtain an estimate of the noise-to-signal ratio using the cross-sectional dispersion of forecasts and compare it to the noise-to-signal ratio implied by the regressions in Section 2.2. First, for each macroeconomic variable, I use the variance of the residuals from the estimated AR(1) process as a proxy for fundamental volatility (σ_v^2). Second, I calculate the average cross-sectional variance of forecasts in the sample, and use (2.9) to obtain an estimate of the signal noise (σ_ω^2). Third, I construct a measure of the noise-to-signal ratio ($\sigma_\omega^2/\sigma_v^2$) by taking the ratio of my "between forecaster" noise measure to the variance of the innovations to the variable from the first step.

I compare this "between forecaster" estimate of the noise-signal ratio to the "within forecaster" estimate that I derived earlier from the forecast error-forecast revision regressions. Figure 2.2 depicts the scatter plot of the two noise-to-signal ratios across the 13 target variables. The two measures are largely estimated from distinct moments of the data, so they represent independent estimates of the same structural parameter. If the structural model was a poor approximation of the expectation formation process, then these two estimates could be very different and unrelated to each other across variables. Although the match is not perfect, there is a clear positive relationship between them. The intercept of the regression line is essentially 0, and one measure explains 90 percent of the variation of the other.

[3] Note that this decomposition result holds only for the consensus forecast, and not for the individual regressions.

[4] The number of usable observations is declining with the forecast horizon, since survey respondents tend to have stronger views about the immediate future. To have the largest possible number of forecasts for estimating the cross-sectional dispersion, I focus on $h = 0$. However, using other horizons does not change the results qualitatively.

Alternative regression specification

So far the analysis emphasized the reduced form result that ex-post forecast errors are predictable using ex-ante forecast revisions both at the aggregate and at individual levels. Under the assumptions of my model, there is an equivalent way to document the same inefficiencies in expectations, which is to look at the correlation between the current forecast revision and past forecast revisions. In Appendix B.1, I show that the following OLS regression can be estimated on consensus forecasts:

$$\overline{z_{t+h|t}} - \overline{z_{t+h|t-1}} = c + \bar{\beta}\left(\overline{z_{t+h|t-1}} - \overline{z_{t+h|t-2}}\right) + \varepsilon_{th} \tag{2.10}$$

with $\bar{\beta} = G\beta$ where β is the regression coefficient in the baseline regression (2.1). Intuitively, the sluggish response of consensus forecasts to new information introduces autocorrelation into forecast revisions. It can also be shown that the same relationship between regression coefficients holds at the individual level. Importantly, this test does not rely on the actual outcomes, and hence it side-steps the issue of what vintage of the actual data should be used when computing forecast errors. For the same reason, equation (2.10) provides an independent way of estimating the deviation from rational expectations that is based purely on the survey responses.

Table 2.6 and 2.7 report the forecast revision regressions on consensus and individual forecasts, respectively. The results overwhelmingly convey the same message as the baseline regressions discussed earlier. At the aggregate level, the majority of regression coefficients is positive and significant, confirming the autocorrelation in the revisions of the consensus forecast. At the individual level, all of the coefficients are significant and their signs conform to the baseline results in each case. Similarly to the baseline regressions, the results for the consensus forecast could be explained by rational models with information rigidities. However, the predictability of individual forecast revisions based on past forecast revisions is a clear violation of rational expectations because it contradicts the law of iterated expectations. Hence, as before, we need to allow for a suboptimal filtering rule to jointly account for the documented facts. Following the same steps as in Section 2.2, I use this alternative specification to uncover the extent of informational inefficiency $(G - G^*)$ in the forecasts of each variable. Again, I find a high degree of consistency in the overall results. As Figure 2.3 demonstrates, the estimated deviations from rationality tend to move together strongly across variables.

2.4 Cross-sectional variation and determinants of forecast inefficiency

The previous sections presented evidence that a simple behavioral model of the expectations formation process goes a long way in explaining otherwise puzzling patterns in the SPF. The model's key feature is a heuristic updating rule in which agents put a weight on new information that is potentially inconsistent with rational expectations. This assumption, however, invites the obvious question of where this behavioral rule comes from. I have already shown that the variation in G across macroeconomic variables seems to be unrelated to the fundamentals that determine the optimal Kalman gain. In this section, I explore the cross-sectional heterogeneity in the efficiency of professional forecasters' expectations. In particular, I ask "Are some forecasters inherently more irrational than others?", and if yes, "Can we relate these differences to observable characteristics?" I document that there is substantial cross-sectional heterogeneity in forecast efficiency for each target variable, and a significant part of this variation is forecaster-specific. I also find that experience and subjective uncertainty are correlated with forecaster-specific deviations from rationality, but they have only limited explanatory power.

Expectations inefficiency across forecasters

So far the analysis focused on pooled estimates from the individual regressions, which assumes that all forecasters are identical. However, forecasters can differ both in the quality of their information and in their ability to use that information efficiently. To shed light on this cross-sectional heterogeneity, I separately reestimate equation (2.6) for each forecaster that has at least 10 valid observations, and I plot the histogram of the β coefficients in Figure 2.4. For most target variables the distribution is fairly symmetric and it is centered around the pooled estimates indicated by vertical lines. This shows that the rejection of rationality based on the pooled regressions were not driven by some influential outliers. In fact, when there is a visible difference between the mode of the forecaster-level β's and the pooled estimates, such as for the AAA yield and the CPI inflation rate, the mode suggests even bigger deviations from rationality. It is also interesting that the degree of cross-sectional heterogeneity is quite similar for all target variables, with the standard deviation of β's ranging between 0.3 and 0.4. As a robustness check, I repeat the exercise with estimates from the alternative regression specification that uses only current and past forecast revisions, and I find qualitatively similar results (Figure 2.5).

The correlation matrix of forecaster-level β coefficients for the 13 target variables reveals that a significant part of the cross-sectional variation in forecast efficiency is forecaster-specific (Table 2.8). Remarkably, all the pairwise correlations are positive, and many of them, especially for the 5 core variables, are significant at the 5% level. This means that a forecaster with relatively strong over-reaction to new information when forecasting one variable (for example, GDP growth) also tends to have relatively strong over-reaction when forecasting another (for example, inflation), and vice versa. Since the forecaster-level β's are estimated on very limited data (often 10-15 observations), they potentially have large measurement error; thus finding this robust pattern of correlations is quite unexpected. Moreover, the estimates of forecast inefficiency from the alternative regression specification paint a very similar picture (Table 2.9). The presence of an individual-specific effect in forecast efficiency across target variables is very interesting, because it suggests that personal characteristics may play an important role in the expectations formation of different agents.

Determinants of forecaster efficiency

I start by constructing a composite measure of the relative informational efficiency of each forecaster's expectations. This measure attempts to extract the common component of the systematic errors that a forecaster makes when forming expectations about different macroeconomic variables. The previous section showed that these errors are positively correlated for all target variables, but the association is most pronounced for the 5 core variables. These variables also have the longest time series and the widest coverage by forecasters, so I constrain my measure to the forecasts of GDP growth, housing starts, industrial production, GDP deflator, and the unemployment rate. I use principle component analysis (PCA) to generate an index for each forecaster that captures most of the cross-sectional variation of the β's associated with these macroeconomic variables.[5] The first principal component accounts for roughly half of the total variance, which is remarkable given the scope of measurement error in the estimated β coefficients. Once again, I also check whether the alternative regression specification yields similar results, and I find that this the case. The forecaster-level indexes extracted from independent estimates of informational efficiency conform to each other quite well in the cross-section as demonstrated in Figure 2.6. This is important further evidence that survey expectations in the SPF are consistent with the simple behavioral model not only *on average*, but also at the *micro-level*.

I also try to relate my measure of relative expectations efficiency to observable characteristics of the forecasters. For example, Malmendier and Nagel (2014) argue that age and personal experiences play an important role in determining consumers' inflation expectations. Unfortunately, the SPF does not record any personal information about the respondents such as age, education or location. However, I can construct two potentially relevant variables from the available survey data. First, I define forecaster i's experience ($EXPERIENCE_i$) as his tenure in the sample, that is the number of quarters between his first and last reported forecast in the SPF. For easier interpretation of the magnitudes, I scale the variable by its standard deviation. Second, I construct a measure of signal noise ($NOISE_i$) from the probability range forecasts that survey respondents submit for the GDP price deflator. Each quarter, forecasters are asked to provide their estimated probabilities that the "annual-average over annual-average" percent change in the GDP price index falls into a number of ranges. Using these range forecasts, I can calculate the standard deviation of forecasters' subjective probability distribution (σ_{it}). The uncertainty captured by this standard deviation reflects both fundamental uncertainty about future innovations to the macro variable (σ_v) and the uncertainty arising from the noisiness of private information (σ_ω). To separate the two, I estimate a fixed effect panel regression:

$$\sigma_{it} = \lambda_i + \delta_t + \varepsilon_{it},$$

where δ_t controls for all sources of uncertainty that are common across forecasters such as macroeconomic shocks. Finally, I equate forecaster-specific information quality with the

[5] Using factor analysis instead of PCA produces numerically almost indistinguishable results.

forecaster fixed effect λ_i. I think of this measure as an empirical proxy for the nosiness of the forecaster's private signal in the model.

Given these measures of experience and signal quality of forecasters, I assess their importance by estimating the following cross-sectional regression:

$$\beta_i^{PC} = c + \gamma_1 NOISE_i + \gamma_2 EXPERIENCE_i + \varepsilon_i, \tag{2.11}$$

where β_i^{PC} is forecaster i's score on the first principle component extracted by the PCA analysis. Our model suggests that $\gamma_1 < 0$. Facing higher signal noise, an agent with rational expectations should put less weight on new information. All else equal, a lower G^* implies that a behavioral agent with a heuristic updating rule is more likely to overreact to his private signals. Since β_i^{PC} is decreasing in the degree of overreaction, we expect a negative coefficient on the signal noise. The interpretation of γ_2 is much less straightforward. It is plausible to hypothesize that more experience improves the informational efficiency of forecasts. Unfortunately, my composite measure on the left-hand side does not capture the absolute size of the deviations from rationality for two reasons. First, the underlying β coefficients in the forecast error-forecast revision regressions have a sign: a negative value signals overreaction to new information, and a positive value signals underreaction. Second, and more importantly, the individual scores in β_i^{PC} provide only a relative ordering of forecasters. A low score indicates that the forecaster tends to overreact to new information compared to the average forecaster, and a high score indicates that the forecaster tends to underreact to new information compared to the average forecaster. As I showed earlier, the average forecaster himself deviates from rational expectations, and thus β_i^{PC} is not directly informative about the degree of irrationality of the individual forecaster.[6] Despite lacking a clear theoretical prediction, I include experience in the baseline regression and I will discuss a potential interpretation of the empirical results.

The results are presented in Table 2.10. The first column shows the simple OLS estimates. The coefficient on signal noise is negative, as predicted by the model, and it is statistically significant at the 5% level. The effect of experience is also significant on average with a positive coefficient. To check whether the results are sensitive to outliers, the second column provides estimates from a robust estimator that automatically detects and down-weights influential outliers. The robust regression yields very similar point estimates and standard errors. To interpret the positive coefficient on experience, I also report quantile regressions for the 0.1 and 0.9 quantiles of the conditional distribution of β_i^{PC}. The results for the 0.1 quantile are very similar to the results from the OLS regressions: higher signal noise pushes the left tale of the distribution down and more experience pushes it upwards. However, the effect of experience gets reversed when looking at the 0.9 quantile: signal noise still has a

[6]This discussion suggests that maybe we should focus on the absolute values of the forecaster-level β's as a measure of expectations inefficiency. However, the correlation of $|\beta_i|$ across different target variables is generally zero in the data. That is, the absolute deviation from rational expectations does not seem to be an inherent property of forecasters – unlike their relative position within the distribution, which is quite consistent across target variables.

negative coefficient, but more experience shifts the right tail of the distribution downwards. Since the left tail of individual β's is typically negative and the right tail is positive (see Figure 2.4), the quantile regressions suggest that the distribution becomes more concentrated near zero as forecaster experience increases.[7] Importantly, these results are consistent with the hypothesis that more experience reduces the deviations from the rational benchmark.

The observed heterogeneity in the efficiency of forecasts is qualitatively consistent with the model's prediction about signal noise and with our intuition about the effect of experience. However, these two observables can account only for a tiny share of the cross-sectional variation; the R-squares range between 3% and 17% in the regressions of Table 2.10. One might conjecture that a larger set of individual characteristics would increase this share substantially. Yet, it is also possible that to a large extent we should think about G as a "deep" type parameter that cannot be derived from other fundamentals.

2.5 Taking stock: Interpretation and implications for macro theory

The key empirical finding of previous sections is that ex-post forecast errors and ex-ante forecast revisions in the SPF are correlated *both* at the aggregate and at the individual level. Broadly speaking, there are two possible interpretations of these reduced form results. On the one hand, it is possible that the survey forecasts do not represent mathematical expected values. If professional forecasters' objective function is different from minimizing the mean squared error, their optimal forecast will deviate from conditional mathematical expectations. For example, I show that strategic motives can lead to forecasters choosing a G different from G^*, which could rationalize the documented informational inefficiencies in the data. On the other hand, if we interpret SPF forecasts as agents' genuine expectations, then my results pose a challenge to the microfoundations of imperfect information models. Specifically, the idea that the aggregate economy responds gradually to shocks *because* a large number of rational agents make Bayesian inferences from noisy signals is not supported. However, this does not imply that the main mechanism embedded in these models is wrong or irrelevant.

Strategic motives in forecasting

As Pesaran and Weale (2006) emphasize, the standard "orthogonality condition" is a valid test of rationality only if forecasters have a loss function that is quadratic in forecast errors, that is if they want to minimize the mean squared error of their forecasts. In various contexts, the literature have considered three types of more general objective functions. First,

[7] In fact, the estimated coefficient on $EXPERIENCE$ is monotonically declining for higher quantiles of the distribution, and it turns negative at about the 0.7 quantile. This also explains why the OLS estimate, which predicts the mean of the conditional distribution, is positive.

an asymmetric loss function can emerge if there are differences in forecasters' costs of over- and underpredicting the target variable. Second, professional forecasters may engage in forecast smoothing if drastic changes in forecasts are hard to explain to their customers. Third, strategic interactions may arise if there is an incentive for forecasters to stay close to or to stay away from the consensus forecast for reputational reasons. Coibion and Gorodnichenko (2015) explicitly consider the standard formulations of asymmetric loss and forecast smoothing, and they show that these modifications cannot explain the behavior of the consensus forecast.[8] Here I demonstrate that a simple model of strategic interactions is not only consistent with the predictability of consensus forecast errors, but it can also account for the informational inefficiency of the reported individual forecasts.

A key feature of the noisy information setup introduced in Section 2.2 is that the cross-sectional variance of priors is always smaller than the uncertainty that any forecaster attaches to his own prior. The intuition for this property is the following. All cross-forecaster dispersion in priors is the result of accumulated past signal noise (ω_{it}), whereas subjective uncertainty about one's own prior is a combination of both signal noise (ω_{it}) and the fundamental volatility of the macro variable (v_t). The relative similarity of priors means that agents can control the expected distance from the average forecast by placing more or less weight on their private signal. Overweighting their private information compared to the Bayesian prescription helps them distinguish themselves from the average, while underweighting private signals allows them to stay close to the consensus forecast in expectation. The rest of the section formalizes this intuition in a static game as in Morris and Shin (2002). Although the game abstracts from the dynamic nature of the original signal extraction problem, it retains all important aspects of the problem and it can be solved in closed form.

Assume that a unit mass of agents want to forecast a macroeconomic variable z. Each agent has the loss function

$$E\left[(1-r)\left(f_i-z\right)^2+r\left(f_i-\bar{f}\right)^2\right], \quad r<1 \tag{2.12}$$

where f_i is agent i's *reported* forecast for variable z, and $\bar{f}=\int_0^1 f_i di$ is the average forecast.

[8] Asymmetric loss over forecast errors is usually modeled by the LINEX loss function (e.g. Capistrán and Timmermann, 2009):

$$L\left(FE_{t+h,t};\phi\right)=\left[\exp\left(\phi FE_{t+h,t}\right)-\phi FE_{t+h,t}-1\right]/\phi^2,$$

where $FE_{t+h,t}=x_{t+h}-x_{t+h|t}$. When $\phi>0$ ($\phi<0$), agents dislike positive (negative) forecast errors more than negative (positive) ones, while $\phi\to 0$ yields the standard mean-squared-error (MSE) objective. Agents engaging in forecast smoothing are usually modeled by the following objective:

$$\min\sum_{j=0}^{h}\gamma^j E_t\left[\left(x_{t+h}-x_{t+h|t+j}\right)^2+\alpha\left(x_{t+h|t+j}-x_{t+h|t+j-1}\right)^2\right],$$

where γ is the discount factor and $\alpha>0$ determines the importance of smoothing. Under certain assumptions, both of these formulations can yield predictability of ex-post forecast errors using ex-ante forecast revisions. However, Coibion and Gorodnichenko (2015) show that both models imply a *negative* regression coefficient, while in the data it is overwhelmingly positive.

The macroeconomic variable evolves according to $z = z_0 + v$, where z_0 is some pre-determined initial value and $v \sim N(0, \sigma_v^2)$ is an unobserved fundamental shock. The timing of the game is the following. Initially, agents start with a diffuse prior about z_0, and then receive a signal $z_i = z_0 + \varepsilon_{0i}$ with $\varepsilon_{0i} \sim iidN(0, \sigma_0^2)$. After this signal, agent i's subjective belief about the initial value is given by $z_0 \sim N(z_i, \sigma_0^2)$, which captures the cumulative effect of all past noise in the dynamic model. Second, nature draws the unobserved fundamental shock v. Now agent i's prior belief about the target variable is $z \sim N\left(z_i, \sigma_0^2 + \sigma_v^2\right)$. Note that the cross-sectional variance of the prior means is only σ_0^2. Third, agents receive a signal about the current state $y_i = z + \omega_i$ where $\omega_i \sim iidN(0, \sigma_\omega^2)$. Finally, agents submit their forecasts. The mathematically correct expectations of agent i is given by

$$E_i\left[z\right] = G^* y_i + \left(1 - G^*\right) z_i \text{ where } G^* = \frac{\sigma_0^2 + \sigma_v^2}{\sigma_0^2 + \sigma_v^2 + \sigma_\omega^2}.$$

However, strategic motives may divert the reported forecasts from conditional expectations.

Following the solution technique of Morris and Shin (2002), I guess and verify that the optimal forecast has the form $f_i = G y_i + \left(1 - G\right) z_i$. Thus, the average forecast can be written as $\bar{f} = Gz + \left(1 - G\right) z_0$. Substituting into the first order condition of minimizing (2.12), we can show that

$$f_i = \left[\left(1 - r + rG\right) G^* + r\left(1 - G\right) b\right] y_i + \left[\left(1 - r + rG\right)\left(1 - G^*\right) + r\left(1 - G\right)\left(1 - b\right)\right] z_i,$$

where $b = \sigma_0^2 / \left(\sigma_0^2 + \sigma_v^2 + \sigma_\omega^2\right)$. This verifies our guess of the optimal forecast as long as

$$\left(1 - r + rG\right) G^* + r\left(1 - G\right) b = G.$$

It is easy to show that the unique solution of this equation is

$$G = \frac{1 - r\frac{\sigma_v^2}{\sigma_0^2 + \sigma_v^2}}{1 - r\frac{\sigma_v^2}{\sigma_0^2 + \sigma_v^2 + \sigma_\omega^2}} G^*. \tag{2.13}$$

If $r = 0$, then forecasters have no strategic incentives and equation (2.13) collapses to the correct mathematical expectations. If $r > 0$ ($r < 0$), then strategic complementarity (substitutability) will induce forecasters to set $G < G^*$ ($G > G^*$), and they will report forecasts that underweight (overweight) their new information compared to the rational expectations benchmark. Note, however, that strategic interaction by itself is not enough to generate systematic inefficiencies in reported forecasts. For example, if information is not noisy ($\sigma_\omega^2 = 0$), every forecaster will report the rational expectations predictions which do not have predictable forecast errors.

Theoretically, the combination of noisy information and strategic motives is able to account for the predictability of consensus and individual forecast errors. Is this mechanism a plausible explanation for the reduced form results of this paper? Two simple observations suggest that other forces are also at play. First, for most macroeconomic variables in the

SPF I find that $G > G^*$ on average, but there are some exceptions. It is hard to see why forecasters would want to distinguish themselves from their peers when predicting one variable but go with the flow when predicting another. Second, previous work has shown that professional forecasters are no worse at forecasting inflation (in terms of mean squared error) than households or financial markets (see Coibion and Gorodnichenko, 2015). Hence, professional forecasters do not seem to underutilize information relative to other agents for whom strategic motives should be less important. On the other hand, strategic motives might help explain the cross-sectional heterogeneity in forecasters' informational inefficiency. If the strength of strategic motives varies across individuals, there will be a forecaster-specific component of informational inefficiency that is fixed across target variables. As I have shown, this prediction holds up in the data.

Inefficient expectations and macro theory

If we interpret SPF forecasts as genuine expectations, then my results provide evidence that a Bayesian view of the expectations formation process is not a good approximation of how agents forecast macroeconomic variables. Moreover, the predictability of individual forecast errors and future revisions using only past revisions of individual forecasts is a completely model-free test of rational expectations. This test does not rely on the true data generating process of the target variable or the exact nature of information constraints. It makes very weak assumptions about information sets, since it only requires that agents have access to their own earlier forecast. Similarly, the test is robust to various forms of heterogeneity considered in the literature such as differences in signal precision and heterogeneity in prior beliefs about long run means. Hence, contrary to the conclusion of recent papers studying consensus survey expectations, the microdata raise important questions about the microfoundations of macroeconomic models where imperfect information and Bayesian updating plays an important role.

Information rigidities are central in many rational expectations theories that tackle some of the most fundamental questions in macroeconomics. One important strand of the literature, for example, attributes the persistent real effects of monetary policy to agents' imperfect and heterogeneous information (Lucas, 1972; Woodford, 2001; Mackowiak and Wiederholt, 2009). In all of these models, some macro variable, such as nominal aggregate demand, follows an exogenous linear stochastic process, agents observe noisy private signals about the current state, and they form optimal Bayesian expectations. With normally distributed innovations, agents' optimal filtering rules are linear. Note that the framework in which I interpreted systematic expectations errors fits this description, except that I did not impose rationality in the filtering rule. And indeed, the data does not support that restriction. Even if the actual innovations were not normal, the linear filtering rules derived in the models would still be the best linear predictors of the true state. Since my empirical evidence against rational expectations is based on linear regressions, forecasters' informational efficiency fails against this weaker benchmark too.

One formal explanation of expectational errors that appear in the literature can be dubbed "signal extraction with wrong parameters." Noisy information models have a somewhat conflicting view of individual decision-makers. On the one hand, they acknowledge the difficulty of information acquisition and the limits of agents' cognitive capacity; but on the other hand they assume that agents have perfect knowledge of the higher level meta-parameters such as the properties of signal noise, and that they can make very complicated inferences from imperfect information. The idea of these papers is that agents may want to calculate mathematical conditional expectations, but they use incorrect parameters when doing so. For example, Gourinchas and Tornell (2004) assume that market participants misperceive the persistence of interest rate innovations, and they show that several exchange rate puzzles can be explained if agents perform rational Bayesian updating with the wrong persistence parameter. Similarly, I could account for my reduced form results in a model where agents misperceive the noise-to-signal ratio and miscalculate the Kalman gain. It is doubtful, however, if this approach can go a long way in explaining expectational errors, since it treats parameter misperception as given. I prefer the more parsimonious approach of treating the updating rule as a rule of thumb which is the primitive of the model, and derive the implications from there.

It is important to emphasize that the failure of agents to use mathematically correct forecasting rules does not, *per se*, undermine the most important aggregate implications of imperfect information macro models. The main predictions of these models, such as the real effects of monetary policy and hump-shaped responses to shocks, do not require that the weight on new information is optimal. They only hinge on agents' partial reaction to signals, which causes information to disseminate slowly at the aggregate level. The heuristic updating rules in this paper's model satisfy this condition, so the aggregate dynamics would remain qualitatively unchanged. Of course, if the updating rule does not respect rational expectations, then some of the comparative statics results of the models might break down. How, if at all, will agents change their information processing after changes in the structure of the economy? Without better understanding the determinants of the heuristic updating rules, we cannot answer such questions.[9]

2.6 Conclusion

This paper revisited the empirical evidence from survey expectations that has been used to support rational expectations models with information frictions. Contrary to the conclusions of recent research based on consensus forecasts, I find that the microdata in the Survey of

[9] It is worth noting a common feature of imperfect information rational expectations models. The comparative statics with respect to the noise-to-signal ratio is often ambiguous or counterintuitive. The main tension is that intuitively noisier information should lead to more volatility in the economy, but at the same time rational agents choose to ignore very noisy signals. In the Lucas model, for example, as the variance of the money supply goes to infinity, real volatility goes to zero. This implication seems to be at odds with our notion about the virtues of predictable monetary policy.

Professional Forecasters are inconsistent with these theories. The main reduced form result is that ex-post forecast errors and forecast revisions are predictable by ex-ante forecast revisions both at the aggregate and at the individual level. These anomalies can be explained in a simple behavioral model of information processing in which agents' updating rules may deviate from the rational benchmark. I show that the survey data are consistent with several cross-sectional and time series predictions of the model. Using three-dimensional panel data on the time path of individual agents' forecasts for different horizons, I estimate the model parameters and quantify the deviations from rational expectations for different target variables. I also find evidence that the degree of informational efficiency is an inherent property of forecasters that carries over from one variable to the other. The paper concludes with potential interpretations of the empirical results, and a discussion of implications for macroeconomic theory.

Table 2.1: Coverage of the Survey of Professional Forecasters

		Number of forecasters		Forecaster tenure (quarters)	
Variable	Sample period	mean	median	mean	median
GDP	1968Q4-2013Q4	38.3	36.0	21.0	15.0
HS		36.9	35.0	20.9	15.0
IP		36.7	35.0	20.4	14.0
PGDP		38.0	35.5	21.2	15.0
UE		38.7	36.5	21.3	15.0
TB3	1981Q3-2013Q4	33.2	33.0	18.3	10.0
AAA		29.3	30.0	17.0	9.0
CPI		33.8	34.0	18.2	11.0
CONS		33.2	34.0	18.6	10.0
FG		31.4	32.0	18.3	10.0
SG		31.4	32.0	18.3	10.0
NRI		32.6	33.0	18.5	10.5
RI		32.5	33.0	18.5	10.0

Notes: The table reports the number of forecasters in a typical quarterly survey (cross-sectional dimension), and the number of quarters that a typical forecaster participates in the survey (time dimension). Respondents typically submit forecasts for several horizons, ranging from the current quarter to four quarters ahead. GDP = real GDP growth rate, HS = housing starts, IP = growth rate of industrial production index, PGDP = inflation rate for GDP deflator, UE = unemployment rate, TB3 = 3 month treasure bill interest rate, AAA = Moody's AAA corporate bond yield, CPI = inflation rate for the consumer price index, CONS = consumption growth rate, FG = growth rate of federal government consumption expenditures, SG = growth rate of state and local government consumption expenditures, NRI = growth rate of non-residential investment, RI = growth rate of residential investment.

Table 2.2: Tests of full-information rational expectations (FIRE) on consensus forecasts

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
Revision	0.35*	0.86***	0.56***	0.73***	0.54***	0.33**	0.08	0.48**	0.02	0.47*	-0.07	0.76***	0.75*
	(0.17)	(0.15)	(0.14)	(0.20)	(0.12)	(0.13)	(0.12)	(0.17)	(0.20)	(0.23)	(0.15)	(0.23)	(0.29)
Constant	-0.09	0.02	-0.68	0.02	-0.03	-0.25***	-0.28***	-0.09	0.31	0.44	0.07	1.14	1.10
	(0.13)	(0.01)	(0.37)	(0.09)	(0.03)	(0.06)	(0.05)	(0.10)	(0.21)	(0.32)	(0.10)	(0.61)	(0.79)
N	689	689	689	689	689	490	490	490	490	490	490	490	490
R2	0.012	0.158	0.035	0.059	0.112	0.049	0.003	0.026	0.000	0.010	0.000	0.035	0.038

Notes: The table reports estimates of equation (2.1). The average ex-post forecast error is regressed on the average ex-ante forecast revision. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.3: Tests of full-information rational expectations (FIRE) on individual forecasts

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
Revision	-0.32***	0.19***	-0.25***	-0.32***	0.24**	0.11	-0.15**	-0.20*	-0.44***	-0.44***	-0.47***	-0.35***	-0.27***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.08)	(0.07)	(0.05)	(0.09)	(0.04)	(0.04)	(0.04)	(0.09)	(0.05)
Constant	-0.32	0.01	-1.11***	0.18	0.00	-0.31***	-0.29***	-0.14	0.25	0.40	-0.05	0.74	-0.08
	(0.16)	(0.01)	(0.32)	(0.13)	(0.03)	(0.06)	(0.05)	(0.12)	(0.14)	(0.32)	(0.12)	(0.64)	(0.96)
N	19382	18613	18504	19199	19657	12551	10921	12643	12623	11829	11858	12336	12355
R2	0.045	0.014	0.022	0.058	0.029	0.007	0.015	0.011	0.077	0.050	0.126	0.032	0.022

Notes: The table reports estimates of equation (2.1) *at the individual level*. Individual forecasters' ex-post forecast errors are regressed on their ex-ante forecast revisions. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.4: Model check: Inclusion of forecaster fixed effects

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
Revision	-0.33***	0.11*	-0.26***	-0.36***	0.19*	0.08	-0.19***	-0.26*	-0.47***	-0.44***	-0.47***	-0.35***	-0.32***
	(0.04)	(0.05)	(0.04)	(0.03)	(0.08)	(0.06)	(0.04)	(0.10)	(0.04)	(0.04)	(0.04)	(0.07)	(0.05)
N	19382	18613	18504	19199	19657	12551	10921	12643	12623	11829	11858	12336	12355
R2	0.050	0.005	0.025	0.082	0.019	0.004	0.027	0.018	0.096	0.049	0.135	0.034	0.032

Notes: The table reports estimates from the same forecast error-forecast revision regression as Table 2.3, but after controlling for forecaster fixed effects. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.5: Model check: Enter current and previous forecast separately

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
$\overline{z_{t+h\|t}}$	0.38*	0.87***	0.58***	0.75***	0.54***	0.30*	0.05	0.33	0.02	0.47*	-0.06	0.79***	0.78**
	(0.17)	(0.16)	(0.14)	(0.20)	(0.12)	(0.13)	(0.11)	(0.19)	(0.20)	(0.19)	(0.15)	(0.23)	(0.30)
$\overline{z_{t+h\|t-1}}$	-0.55*	-0.87***	-0.75***	-0.77***	-0.60***	-0.34*	-0.13	-0.65***	-0.02	0.21	0.39	-0.84**	-0.58*
	(0.22)	(0.17)	(0.20)	(0.20)	(0.12)	(0.14)	(0.11)	(0.18)	(0.25)	(0.27)	(0.22)	(0.29)	(0.26)
Constant	0.41	0.03	-0.07	0.13	0.30*	-0.02	0.32	0.95***	0.22	-0.52	-0.50*	1.38	0.56
	(0.32)	(0.04)	(0.49)	(0.14)	(0.12)	(0.12)	(0.18)	(0.25)	(0.38)	(0.33)	(0.23)	(0.86)	(0.64)
$p(\beta_1 + \beta_2 = 0)$	0.0793	0.8440	0.2026	0.4726	0.0058	0.0779	0.0024	0.0000	0.8041	0.0000	0.0034	0.6985	0.2439
N	689	689	689	689	689	490	490	490	490	490	490	490	490
R2	0.018	0.158	0.040	0.061	0.135	0.073	0.086	0.075	0.013	0.046	0.016	0.035	0.048

Notes: The table reports estimates of equation (2.8). The ex-post consensus forecast error is regressed on the current and the previous consensus forecast. The p-value for $H_0 : \beta_1 + \beta_2 = 0$ is also provided in the table. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.6: Model check: Alternative specification - Consensus forecasts

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
Revision(-1)	0.33***	0.56***	0.40***	0.29*	0.43***	0.30*	0.24***	0.24	-0.03	-0.01	-0.16	0.50***	0.39***
	(0.09)	(0.10)	(0.11)	(0.11)	(0.07)	(0.12)	(0.07)	(0.18)	(0.12)	(0.17)	(0.09)	(0.13)	(0.11)
Constant	-0.19***	-0.00	-0.37***	0.00	0.01	-0.10*	-0.09*	-0.08	-0.08	0.17	-0.04	-0.11	-0.39
	(0.05)	(0.01)	(0.10)	(0.04)	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)	(0.13)	(0.04)	(0.15)	(0.24)
N	514	514	514	514	514	366	366	366	366	366	366	366	366
R2	0.062	0.276	0.074	0.058	0.206	0.106	0.059	0.017	0.001	0.000	0.019	0.146	0.107

Notes: The table reports estimates of equation (2.10). The average forecast revision is regressed on the previous average forecast revision. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.7: Model check: Alternative specification - Individual forecasts

	(1) GDP	(2) HS	(3) IP	(4) PGDP	(5) UE	(6) TB3	(7) AAA	(8) CPI	(9) CONS	(10) FG	(11) SG	(12) NRI	(13) RI
Revision(-1)	-0.25***	0.06	-0.23***	-0.29***	0.17**	0.12*	-0.07	-0.23***	-0.32***	-0.34***	-0.31***	-0.22***	-0.19***
	(0.03)	(0.04)	(0.03)	(0.03)	(0.06)	(0.05)	(0.03)	(0.06)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)
Constant	-0.26***	-0.01	-0.50***	0.05	0.03	-0.16***	-0.12***	-0.12*	-0.11*	0.21	-0.08*	-0.39	-1.17***
	(0.07)	(0.01)	(0.12)	(0.05)	(0.02)	(0.04)	(0.04)	(0.05)	(0.05)	(0.12)	(0.04)	(0.21)	(0.33)
N	11468	10944	10883	11308	11651	7659	6631	7689	7776	7243	7258	7585	7599
R2	0.064	0.003	0.044	0.092	0.034	0.017	0.005	0.031	0.083	0.089	0.089	0.040	0.031

Notes: The table reports estimates of equation (2.10) *at the individual level*. Individual forecasters' forecast revisions are regressed on their previous forecast revisions. For the explanation of macroeconomic variables in the header, see the notes under Table 2.1. Driscoll-Kraay (1998) standard errors are in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.8: Correlation matrix of individual β's from the baseline specification

	GDP	HS	IP	PGDP	UE	TB3	AAA	CPI	CONS	FG	SG	NRI	RI
GDP	1.00												
HS	0.29^{***}	1.00											
IP	0.26^{***}	0.34^{***}	1.00										
PGDP	0.35^{***}	0.16^{*}	0.14^{*}	1.00									
UE	0.30^{***}	0.38^{***}	0.36^{***}	0.03	1.00								
TB3	0.05	0.18^{*}	0.18^{*}	0.05	0.33^{***}	1.00							
AAA	0.04	0.03	0.09	0.04	0.00	0.25^{**}	1.00						
CPI	0.16	0.26^{**}	0.09	0.29^{***}	0.08	0.00	0.01	1.00					
CONS	0.10	0.20^{*}	0.03	0.04	0.02	0.10	0.05	0.07	1.00				
FG	0.01	0.20^{*}	0.14	0.05	0.06	0.05	0.01	0.01	0.05	1.00			
SG	0.03	0.02	0.10	0.01	0.05	0.01	0.03	0.01	0.03	0.02	1.00		
NRI	0.31^{***}	0.23^{**}	0.37^{***}	0.10	0.37^{***}	0.17	0.02	0.07	0.07	0.19^{*}	0.01	1.00	
RI	0.16	0.21^{*}	0.31^{***}	0.08	0.25^{**}	0.07	0.11	0.04	0.24^{**}	0.02	0.10	0.35^{***}	1.00

Notes: The table reports pairwise correlations of the individual β coefficients in the baseline specification estimated for different macroeconomic variables. For each forecaster i and macroeconomic variable k , I regress forecast errors on forecast revisions, and I save the regression coefficient β_{ik}. The element of the matrix at variable k_1 and k_2 gives $Corr_i(\beta_{ik_1}, \beta_{ik_2})$. For the explanation of macroeconomic variables, see the notes under Table 2.1. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.9: Correlation matrix of individual β's from the alternative specification

	GDP	HS	IP	PGDP	UE	TB3	AAA	CPI	CONS	FG	SG	NRI	RI
GDP	1.00												
HS	0.13	1.00											
IP	0.29***	0.32***	1.00										
PGDP	0.35***	0.15*	0.14*	1.00									
UE	0.14*	0.36***	0.16*	0.13	1.00								
TB3	0.25**	0.13	0.14	0.04	0.25**	1.00							
AAA	0.09	0.09	0.04	-0.01	0.13	0.07	1.00						
CPI	-0.03	0.23**	0.01	0.13	0.14	0.19*	0.08	1.00					
CONS	0.33***	-0.06	0.16	0.06	0.14	0.35***	-0.04	0.19*	1.00				
FG	0.05	0.10	0.04	0.08	-0.03	0.13	0.06	0.19*	0.04	1.00			
SG	-0.04	0.02	0.09	-0.06	0.01	0.11	0.28**	-0.07	-0.05	-0.01	1.00		
NRI	0.32***	0.10	0.16	0.15	-0.04	0.06	0.09	0.09	0.06	0.26**	-0.05	1.00	
RI	0.12	0.36***	0.31***	0.12	0.24**	0.20*	0.05	0.06	0.12	-0.04	0.39***	0.05	1.00

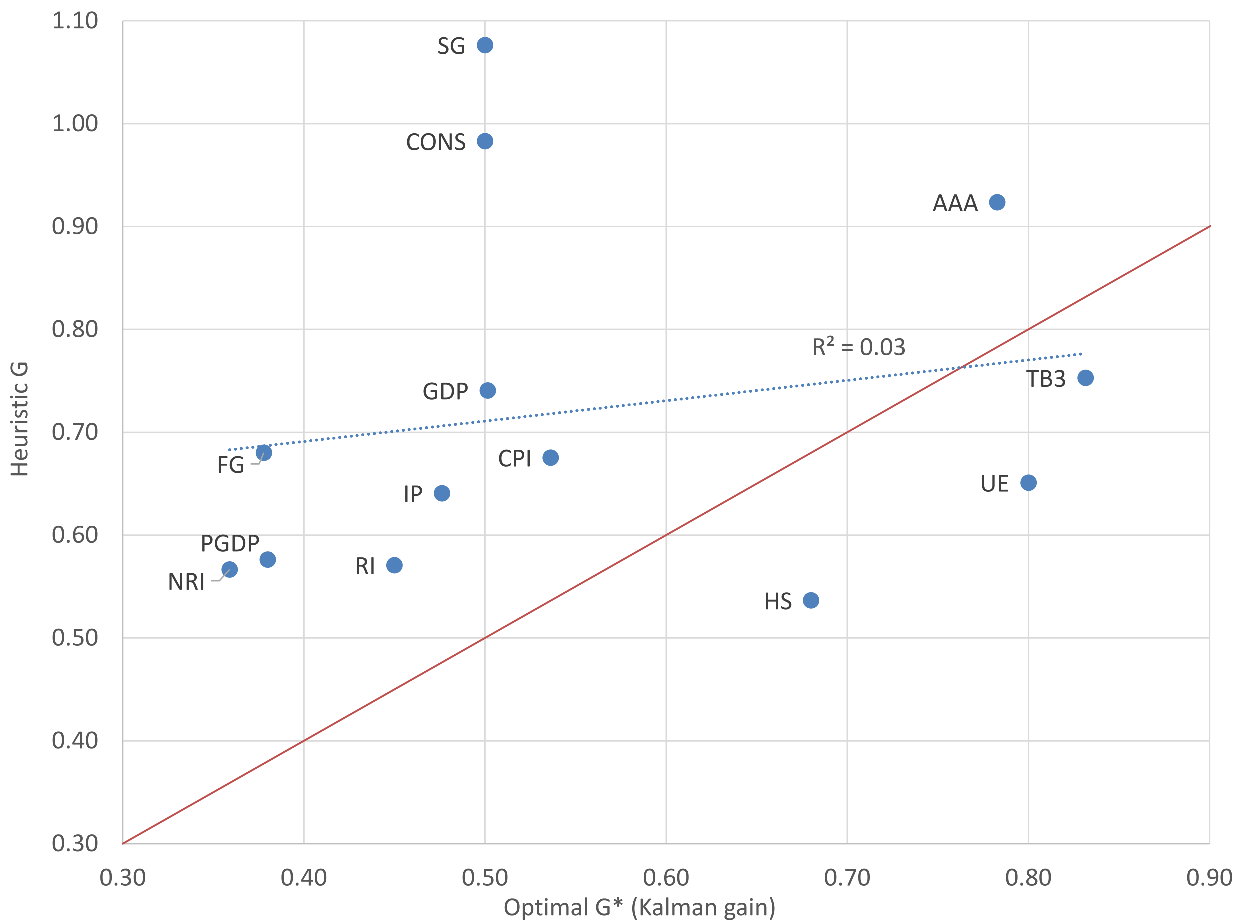
Notes: The table reports pairwise correlations of the individual β coefficients in the alternative specification estimated for different macroeconomic variables. For each forecaster i and macroeconomic variable k , I regress forecast revisions on lagged forecast revisions, and I save the regression coefficient $\bar{\beta}_{ik}$. The element of the matrix at variable k_1 and k_2 gives $Corr_i(\bar{\beta}_{ik_1}, \bar{\beta}_{ik_2})$. For the explanation of macroeconomic variables, see the notes under Table 2.1. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 2.10: Cross-sectional determinants of forecast efficiency

	(1) OLS	(2) Robust reg.	(3) Quantile: D1	(4) Quantile: D9
NOISE	-0.75*	-0.90*	-1.16**	-0.40
	(0.37)	(0.36)	(0.42)	(0.22)
EXPERIENCE	0.24*	0.29**	0.42***	-0.37*
	(0.11)	(0.10)	(0.12)	(0.18)
Constant	-0.24	-0.43*	-1.99***	2.66***
	(0.21)	(0.20)	(0.23)	(0.54)
N	176	176	176	176
(pseudo-) R2	0.055	0.066	0.178	0.032

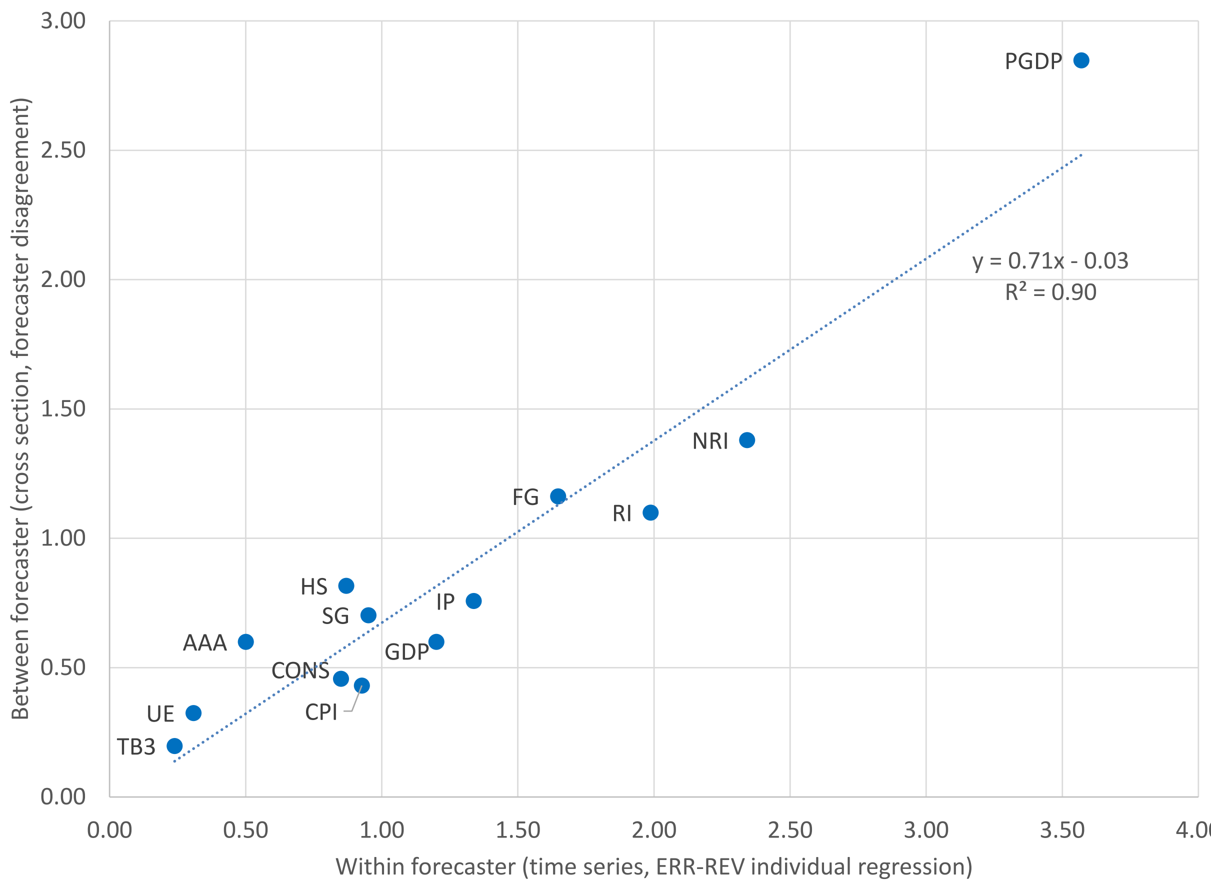
Notes: The table reports estimates of equation (2.11) in the text. The dependent variable is the forecaster-level index of relative expectations efficiency (β_i^{PC}). NOISE is a proxy for the noise in agent i's signal and EXPERIENCE is the forecaster's normalized tenure as defined in the text. Column (1) is the standard OLS estimate. Column (2) is estimated by an iterative regression procedure which detects and down-weights outliers. Column (3) and (4) present quantile regressions: D1 is the 0.1 quantile and D9 is the 0.9 quantile. For the quantile regressions the pseudo-R2 is reported. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Figure 2.1: Comparing the heuristic and the optimal updating rule



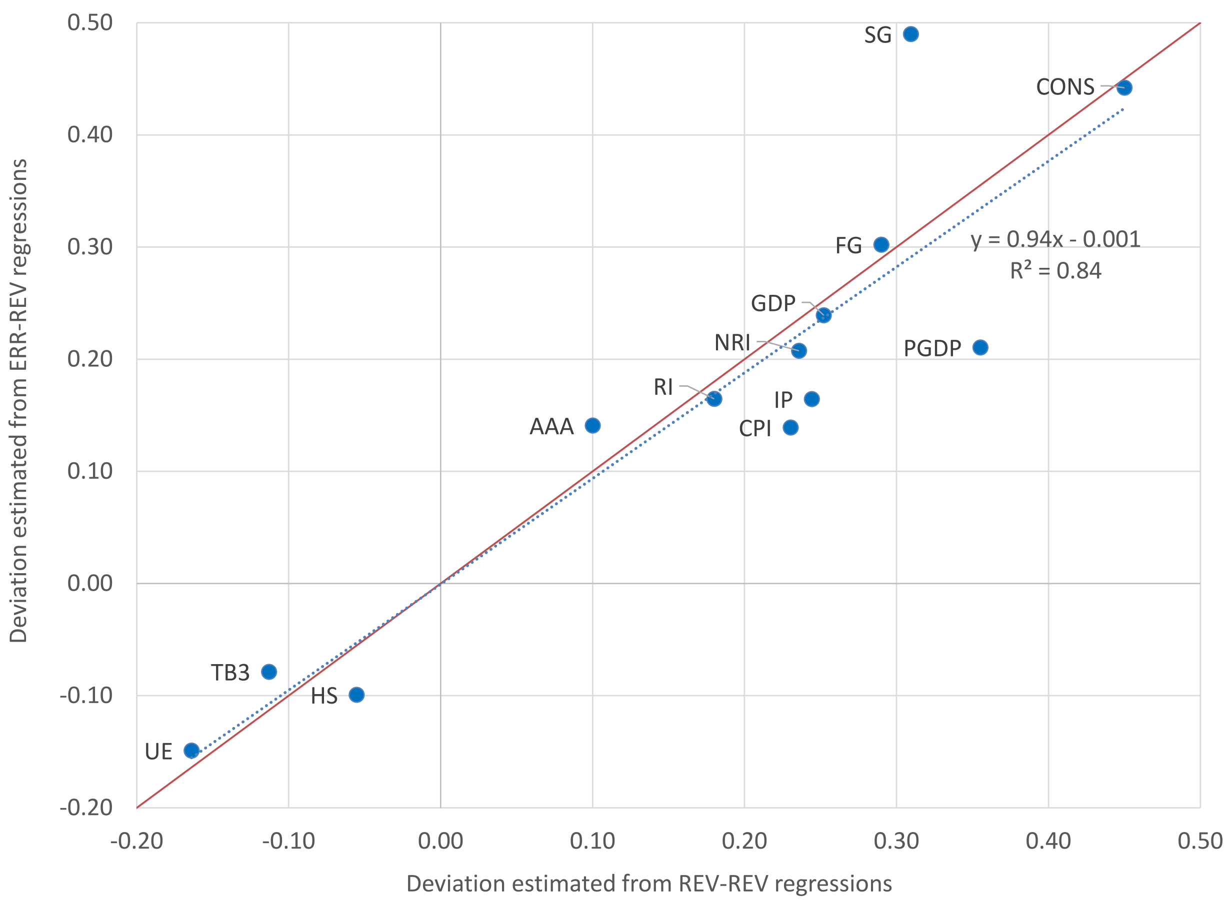
Notes: The figure plots the weight on new information in the heuristic updating rule (vertical axis) and the weight that a Bayesian agent would use (horizontal axis) for the 13 variables in the SPF. Both parameters are calculated from the reduced form regressions using the structure of the model as explained in Section 2.2. For the explanation of macroeconomic variables, see the notes under Table 2.1.

Figure 2.2: Model check: Implied noise-to-signal ratios



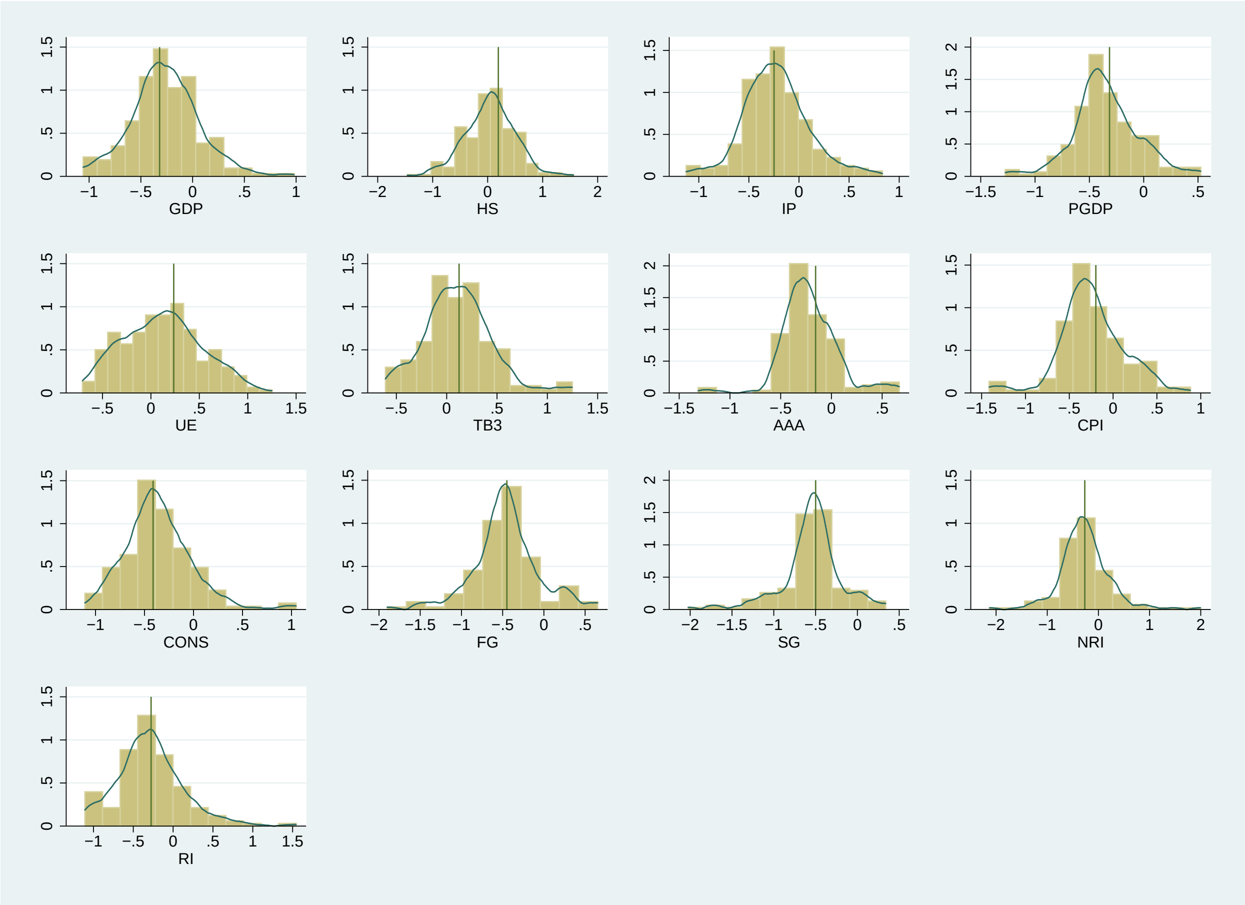
Notes: The figure plots the noise-to-signal ratio implied by the baseline regressions (horizontal axis) and the noise-to-signal ratio implied by forecaster disagreement (vertical axis). The derivation of signal noise from the cross-sectional dispersion of forecasts is explained in Section 2.3. For the explanation of macroeconomic variables, see the notes under Table 2.1.

Figure 2.3: Model check: Deviations from the rational expectations ($G - G^{*}$)



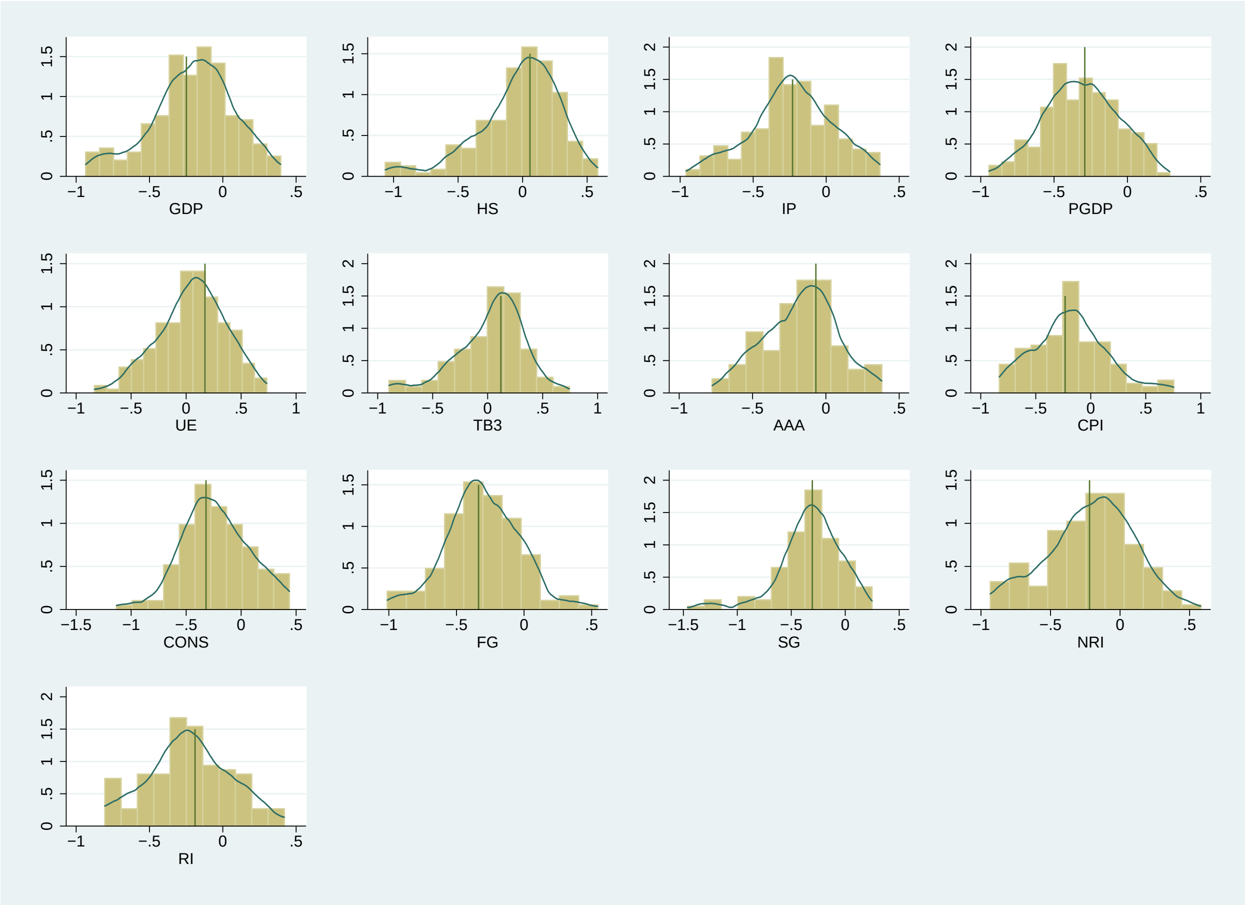
Notes: The figure plots the deviations from the rational expectations benchmark for the 13 variables in the SPF. The vertical axis shows the estimates from the baseline specification which uses the actual realization of the data, while the horizontal axis depicts the estimates from the alternative specification that relies only on the forecast data. For the explanation of macroeconomic variables, see the notes under Table 2.1.

Figure 2.4: Cross-sectional heterogeneity in forecast efficiency (baseline specification)



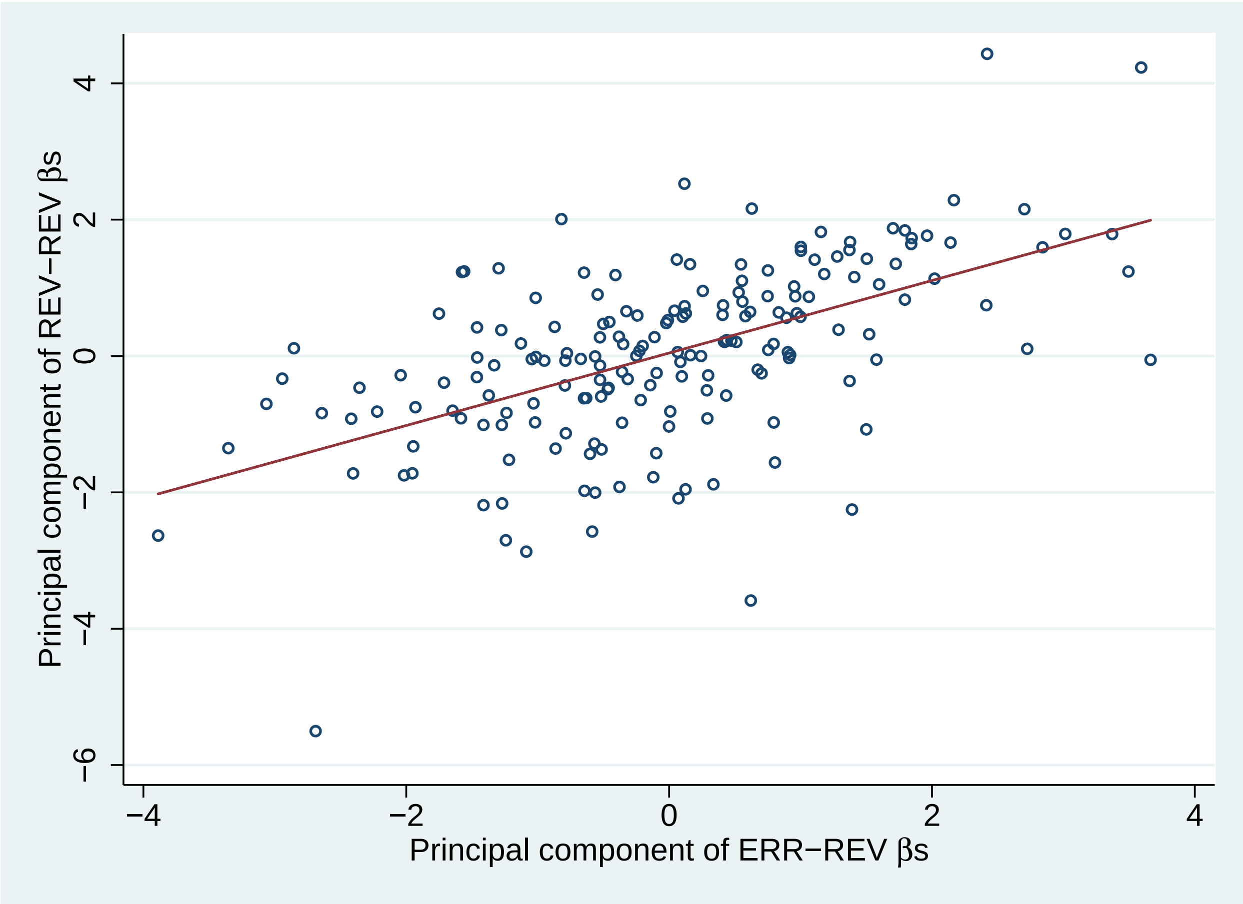
Notes: The figure plots the distribution of individual β coefficients in the baseline specification. For each forecaster i and macroeconomic variable k , I regress forecast errors on forecast revisions, and I save the regression coefficient β_{ik}. Each cell shows the histogram and estimated kernel density of β_{ik} for a particular variable k. The top and bottom 1% of the βs are excluded as outliers. For the explanation of macroeconomic variables, see the notes under Table 2.1.

Figure 2.5: Cross-sectional heterogeneity in forecast efficiency (alternative specification)



Notes: The figure plots the distribution of individual β coefficients in the alternative specification. For each forecaster i and macroeconomic variable k , I regress forecast revisions on lagged forecast revisions, and I save the regression coefficient $\bar{\beta}_{ik}$. Each cell shows the histogram and estimated kernel density of β_{ik} for a particular variable k. The top and bottom 1% of the βs are excluded as outliers. For the explanation of macroeconomic variables, see the notes under Table 2.1.

Figure 2.6: Model check: Forecaster-level informational efficiency



Notes: The figure plots scores on the first principal component which was extracted from individual forecasters' efficiency for several macroeconomic variables. On the horizontal axis, forecast efficiency is measured by the individual β coefficients from the baseline regression of forecast errors on forecast revisions. On the vertical axis, forecast efficiency is measured by the individual β coefficients from the alternative regression of current forecast revisions on lagged forecast revisions. The two composite indexes of relative forecast efficiency are highly correlated (r=0.65).

Chapter 3

Bank Competition, Leverage, and the Introduction of the Euro

3.1 Introduction

What drives capital structure decisions in the banking sector? Without convincingly answering this question, our chances of designing a welfare improving regulatory system are minimal. Many policymakers and academic researchers view the buildup of leverage in the financial sector as an essential ingredient for the severity of the recent financial crisis. Nevertheless, we do not have a good understanding how the individually optimal decisions of financial intermediaries can lead to such suboptimal outcome at the macro level. Why did banks not internalize the costs of a potential financial meltdown whose probability increased every time they opted for more debt financing? As one possible answer to this question, this paper revisits the proposition that in the banking sector there is a trade-off between competition and systemic stability.[1]

In the aftermath of the global financial crisis there has been an intense debate both in policy making institutions and the academic community about how the deficient regulatory frameworks need to be reformed (e.g. Bernanke, 2011; Blinder, 2010; Basel Committee on Banking Supervision, 2010; U.S. Treasury, 2009). A widely shared view among all these commentators is that the new approach must involve substantially higher capital requirements on financial firms, particularly in good times. It is argued that increased capital requirements offer a clear benefit in terms of increased systemic stability. But will these higher capital requirements lead to increased costs for borrowers?

A pervasive view expressed by representatives of the banking industry is that equity is very "expensive", and that equity requirements impose large costs on the financial system and on the economy. Bankers have mounted a campaign against increasing equity require-

[1] In the paper, I use the words bank, financial firm, and financial intermediary interchangeably. What matters for the argument is the function of the entity, and not its legal form. Similarly, although the empirical analysis focuses on traditional banks, the lessons are equally relevant for the "shadow banking system."

ments. Policy makers and regulators are particularly concerned by assertions that increased equity requirements would increase borrowing costs, restrict bank lending and would impede economic growth.[2]

However, more careful recent academic research has shown that these passionate statements are somewhat exaggerated, at least quantitatively.[3] From a theoretical point of view, this literature maintains that any discussion of capital structure must start from the Modigliani-Miller paradigm, and consider possible deviations from this idealized world. Three main deviations were suggested: the tax deductibility of interest payments, a convenience premium on "money-like" short-term bank debt and asymmetric information problems a la Myers and Majluf (1984). These papers conclude that the effects should be quantitatively small in the banking sector. Regarding the tax advantage of debt, it is also important to point out that this is not a social cost, but rather a distortion in the tax code itself.

As an empirical matter, one can also examine the historical record to see if there is a quantitatively meaningful relationship between equity ratios and macroeconomic outcomes. One important observation is that U.S. banks operated with much higher equity ratios in the past – above 50% in the 1840s, about 25% at the end of the 19th century and about 15% in the first decades of the 20th century.[4] A similar trend is apparent in the UK. However, these large fluctuations in capital ratios did not translate into systematic movements in credit availability or economic activity.

Given the evidence discussed so far, it is hard to understand why banks are so inclined to use debt financing instead of equity, and why they employ armies of lobbyists to fight higher capital requirements. In this paper, I make the case that increasing competition in the banking sector might be a good place to look at for answers. My argument builds heavily on the informal "competition hypothesis" of Hanson, Kashyap and Stein (2011, henceforth HKS), but also complements it in two important ways.

First, I develop a simple model in which debt financing is cheaper than equity due to the tax shield it offers. Banks have an incentive to stay in business because they have market power and are able to earn economic profits in the next period; i.e. the bank charter value is positive. Managers, acting in the interest of shareholders, trade-off the tax benefits of debt financing against the increased probability of default which would result in the loss of charter value. As competition increases, the charter becomes less valuable, and banks optimally choose higher leverage.

Second, as an empirical test of the theory, I examine the evolution of equity ratios in the European banking sector around the introduction of the euro. I present some narrative evidence that the creation of the single currency was perceived as a substantial shock to competition in European banking. I also describe the mechanisms through which competitive

[2]See Admati et al. (2013) for a collection of interesting quotes representing this view.

[3]Admati et al. (2013) give an exhaustive refutation of various claims suggesting that equity is expensive *per se*. See Kashyap, Stein and Hanson (2010) for a similar argument and empirical evidence from the US. Miles, Yang and Marcheggiano (2013) use UK data, and estimate the optimal level of bank equity which is much higher than banks have used in recent years.

[4]This number was in the range of 7-8% on average before the crises.

pressure emerged. A panel regression analysis provides evidence which is consistent with the competition hypothesis: banks in countries that entered the currency area tended to tilt their capital structure more towards debt than their peers in non-euro countries.

The paper is organized as follows. Section 3.2 introduces the informal version of the competition hypothesis from HKS (2011). Section 3.3 presents the model. Section 3.4 describes the perceived and actual effects of the euro on competition in the eurozone banking sector. Section 3.5discusses the data and methodology, and section 3.6 contains the main empirical results. Section 3.7 does some robustness checks. Section 3.8 concludes. Appendix C.1 describes the data and related issues in greater detail and Appendix C.2 presents some aspects of the model which are not worked out in the text.

3.2 Competition hypothesis: an informal exposition

The unanimous conclusion of various recent studies discussed in the introduction is that even substantially higher capital requirements would only marginally increase banks' funding costs. If this is true, however, then why do banks nowadays feel compelled to operate in such a highly leveraged fashion? Notice that all the arguments for equity being an expensive form of financing also apply to nonfinancial firms. Yet, we never observe nonfinancial firms with leverage ratios typical in the banking sector.

HKS (2011) offer an explanation to this seeming paradox, which they call the competition hypothesis. They argue that

> "[...] the resolution of this puzzle has to do with the nature of competition in financial services. The most important competitive edge that banks bring to bear for many types of transactions is the ability to fund themselves cheaply. Thus, if Bank A adopts a capital structure that raises its cost of funding relative to other intermediaries by 20 basis points, it may lose most of its business, (or become much less profitable, since the return on assets in banking is on the order of 125 basis points). Contrast this with, say, the auto industry, where cheap financing is only one of many possible sources of advantage: a strong brand, quality engineering and customer service, and control over labor costs may all be vastly more important than a 20 basis-point difference in the cost of capital." (p. 20).

One suggestive piece of evidence for this competition hypothesis comes from the comparison of equity ratios of small and large U.S. banks. The authors show that there is a strong inverse relationship between bank size and capital ratios, with small banks having capital ratios more than double those of large banks. The ability of small banks to operate at higher capital levels probably reflects the softer degree of competition in their core line of business which is informationally intensive "relationship lending" that creates specificity between firms and their lenders (Rajan, 1992). A more formal test of the hypothesis makes use of changes in U.S. state branching and interstate banking regulations that exogenously altered the competitive environment in the affected states.[5] A panel regression analysis reveals that a regulatory shock that increased the degree of competition in a given state led to a significant decline in the average equity ratio of banks in that state.

The validity of the competition hypothesis has important policy implications. If it is an adequate description of at least part of the story, substantially stricter capital regulation of financial firms is appealing. It would seem to hold the promise of reducing competition on a dimension that creates systemic risk, while at the same time not raising loan rates by much. However, the argument presented by HKS is not completely satisfying from a theoretical point of view. Furthermore, their empirical evidence is limited to the US, while the problem of high leverage in the financial sector is worldwide as the global crisis has shown. In the following section I will try to complement the HKS argument along these two lines.

First, I develop a simple model which, although extremely stylized, helps clarify some subtle points in the theory. Since HKS does not present or refer to any formal model, it is not easy to grasp the exact mechanism they have in mind. In particular, if equity is more expensive than debt, why do not banks use debt exclusively? And if there is some privately optimal capital structure, what is the mechanism that causes banks to weigh more heavily towards debt when competition increases? In my simple framework banks trade-off the tax benefits of debt financing against the increased probability of default; the latter decreases expected future economic profits which in turn depend on the level of competition.

Second, as a contribution on the empirical side, I use the introduction of the euro as a shock to competition in the European banking sector, and try to detect changes in equity ratios. I will discuss several reasons that establishing a currency union enhances competition in the banking sector, and show that it was an expected effect of the euro at the time. This new evidence from Europe might complement our view of the empirical validity of the competition hypothesis.

3.3 Competition hypothesis: a simple model

The purpose of this section is to demonstrate formally the main trade-off between debt and equity, and how it is affected by the degree of competition. Consider a highly stylized static model, where at the beginning of the period a bank faces an exogenous investment opportunity, and has to decide about its capital structure; i.e. how much of the investment is financed by debt versus equity. The capital structure decision is the only one made by the bank. The bank starts the period with some cash, C, on its asset side, and only equity on the liabilities side (E_{old}). The number of outstanding shares is s, and the bank can obtain new equity (E_{new}) by issuing new shares. The managers are acting in the interest of the original shareholders, and everybody in the economy is risk neutral. The price of the investment is A, and it has a stochastic gross return R which is realized at the end of the period. The investment can be interpreted as a loan portfolio that the bank can extend to its customers.

[5] Prior to 1970, the majority of U.S. states had restrictions on intrastate branching and allowed absolutely no interstate banking. These restrictions were abolished at different times by different states.

The risk free interest rate is normalized to zero. I assume that the bank does not have enough cash to finance the whole investment project, so it has to raise new funds from the market equal to $A - C > 0$.[6]

Let us denote the face value of the issued debt by D, and its market value at the beginning of the period by Z. The two can differ, since debt holders do not receive the full promised amount in case the bank defaults. The number of newly issued shares is denoted by s'. The total payoff of debt holders and equity holders at the end of the period is V_D and V_E, respectively. An illustrative balance sheet of the model's representative bank is depicted in Figure 3.1. With this notation, we can write the bank's optimization problem as follows:

$$\max_{s',D} \frac{s}{s+s'} E\left[V_E\right]$$

$$\text{s.t.} \quad \frac{s'}{s+s'} E\left[V_E\right] = A - C - Z \qquad (\mathrm{PC_E})$$

$$E\left[V_D\right] = Z \qquad (\mathrm{PC_D})$$

$$0 \leq Z \leq A - C$$

$$s' \geq 0$$

The interpretation of the equations is straightforward. The bank wants to maximize the expected value of existing shareholders' equity, but faces two participation constraints. $\mathrm{PC_E}$ and $\mathrm{PC_D}$ simply state that our risk neutral financiers break even. Using $\mathrm{PC_E}$, we can express s' as

$$s' = s \frac{A - C - Z}{E\left[V_E\right] - (A - C - Z)}.$$

Using this expression and $\mathrm{PC_D}$, we re-write the objective function as follows:

$$\frac{s}{s+s'} E\left[V_E\right] = E\left[V_E\right] + E\left[V_D\right] - (A - C) \qquad (3.1)$$

This result reflects the fact that in the absence of any incentive problems the managers, acting in the interest of original shareholders, maximize firm value, i.e. the total expected payoff of debt and equity holders. However, I did not specify yet how these payoffs are determined. This is where I introduce two reduced form frictions which will cause deviations from the Modigliani-Miller assumptions.

First, in line with the view expressed by representatives of the banking industry, debt will be a "cheaper" source of financing than equity. I introduce this friction through the tax advantage of debt, and I model it by a transfer from the government which is proportional

[6] I assume that the bank invests all its cash into the risky asset. The only implication of this is that I do not have to carry the constant C all over the derivations.

to the face value of issued debt.[7] That is, if the bank is able to honor its outstanding debt D at the end of the period, it receives a transfer γD from the government.[8] Second, following Keeley (1990), I assume that banks have market power and they are able to earn monopoly rents. This implies that the value of staying in business, i.e. the bank's charter value is positive.[9] The charter value is inherently related to the degree of competition in the banking sector; sharper competition reduces monopoly rents, and hence the charter value.

Let us denote the charter value by X. At the end of the period, if the bank is insolvent (that is, its assets, not including the charter value, are less than debt obligations), equity holders receive nothing, debt holders receive whatever is left in the bank and the charter is lost. If the bank is solvent, however, the bank retains its charter value and continues to operate for another period.

Taking into account these two frictions, the expected payoffs of equity and debt holders, respectively, are

$$\begin{aligned}
E\left[V_E\right] &= \Pr\left[RA > D\right]\Big\{E\big[RA \mid RA > D\big] - (1-\gamma)\,D + X\Big\},\\
E\left[V_D\right] &= \Pr\left[RA > D\right]D\\
&\quad + \Pr\left[RA \leq D\right]E\big[RA \mid RA \leq D\big].
\end{aligned}$$

Substituting these into the bank's objective function (3.1), we obtain

$$E\left[V_E\right] + E\left[V_D\right] - (A - C) = E\big[RA\big] + \Pr\left[RA > D\right](\gamma D + X) - (A - C)$$

Dropping constant terms, and introducing the notation $d = D/A$ and $x = X/A$, we can re-write the bank's optimization problem with the debt ratio as the choice variable:

$$\max_{d\in[0,d_{\max}]} \Pr\left[R > d\right](\gamma d + x), \tag{3.2}$$

where d_{max} is the maximum level of d allowed by the constraints.[10] For now, let us assume that the distribution of the random variable R and the parameters are such that we have an

[7] There are many ways to motivate the preference for debt. One example is (mispriced) deposit insurance: if depositors are insured by the government, they require lower interest rates (Keeley, 1990). Notice that this does not explain the enormous increase in wholesale debt financing on banks' balance sheets in recent years. Another way is to introduce informational asymmetries: managers have an incentive to issue equity when they think the public is overpricing their company, which leads to a lemons problem (Myers and Majluf, 1984). Although arguably this is the most convincing theory, these models are usually very involved technically. Since the underlying cause of the advantage of debt is not important for my argument, I use the simplest way to model it.

[8] In reality, the tax benefit of debt comes from the tax deductibility of interest payment. This means that the real world counterpart of γ is the product of the interest rate and the corporate tax rate.

[9] In a more realistic dynamic model the charter value would obtain as the discounted present value of future rents.

[10] This value depends on the distribution of R, since d is defined with the face value of debt (D) and not by its market value at the beginning of the period (Z). Also, notice that in the empirical part I will focus on the equity ratio and not the debt ratio. In the model, it is more straightforward to write the optimization on terms of d, but we could equivalently express everything in terms of the equity ratio.

interior optimum with $0 < d < d_{max}$. (In Appendix C.2, I elaborate on the conditions under which we get meaningful comparative statics, and show that for sensible distributions the following intuition is correct.)

Equation (3.2) demonstrates most transparently my interpretation of the competition hypothesis of HKS. Owing to the unique nature of the financial services industry, probably the most important decision a bank has to make concerns its average funding costs. In my stylized model this is the only choice variable of the bank, represented by d. The basic tension lies between reaping the tax benefits of debt and losing the charter value which is affected by the degree of competition. Equation ((3.2) suggests the following intuitive results:

1. If $\gamma = x = 0$, there is no optimal capital structure that maximizes shareholder value. In the absence of tax shield and monopoly power, there is no tension between equity and debt. This corresponds to the Modigliani-Miller world, in which capital structure is irrelevant.

2. If $\gamma > 0$ and $x = 0$, there is an optimal debt ratio which maximizes shareholder value. Managers want to benefit from the tax shield, but they realize that increasing debt increases the probability of default. Notice that even without positive charter value the bank has an incentive to restrain from full debt financing, since in default states the tax shield is worthless.

3. The most interesting implication of the model is that the charter value, x, affects the trade-off discussed in point 2. With higher charter value, the loss in default is bigger, so the second term in the product becomes more important. For sensible return distributions, this implies that managers will put more weight on staying in business, and will reduce leverage (see Appendix C.2).

Figure 3.2 uses numerical examples to illustrate the model's mechanism. In the left panel, I assumed that R has a lognormal distribution, while the right panel assumes a uniform return distribution. I plotted the bank's objective function for three levels of the charter value: high, medium and low. One can interpret these as corresponding to mild, medium and sharp competition in the banking sector. The graphs confirm our intuition. As competition increases, the bank's charter value shrinks, and the optimal level of leverage rises.

After developing intuition in a stylized model, the remaining part of the paper carries out an empirical test of the theory. The quasi-experiment I consider is the introduction of the euro.

3.4 The effect of the euro on competition in the banking sector

Competition policy in the EU

As the following section explains in detail, my identification strategy rests on the assumption that the euro was a one-time massive shock to bank competition in the affected countries. However, if I want to use the timing of euro introduction as exogenous variation in competition, I have to make sure that there was no other significant contemporaneous change in relevant policies which could confound the detected effect.[11]

Carletti and Vives (2009) and Casu and Girardone (2009) give extensive reviews of the evolution of regulation and competition policy in the European banking sector. Based on these studies, Table 3.1 shows the most important regulatory measures impacting on the EU banking and financial sectors. The Second Banking Directive, introduced in 1989, is doubtlessly the most important of the EU directives on banking regulation. It created a single banking license throughout the EU, and regulated the operation of banks based in one member state in other EU countries The direction also stipulated common capital adequacy requirements for banks in member states.

Although the Second Banking Directive established one single banking market in theory, some details were not sufficiently worked out, which allowed governments to exercise some degree of protection of their national banking systems. EU legislation has been trying to gradually fill these gaps. The rapid development of the financial industry also imposed new challenges, which led to several subsequent directives on specific issues, such as investment services and e-money.

My reading of the literature suggests that EU regulation and competition policy in the banking sector have been a slow moving process, with a continuous cat-and-mouse game between national governments and EU policymakers. There was certainly no one-time shock comparable in magnitude to the introduction of the euro.[12] This gives me some confidence that the effects I find are caused by the competition enhancing impact of the euro, and not by some other policy change. The channels through which the euro might affect competition are discussed in the next section.

The effect of the euro: channels and contemporary expectations

In this paper, I will not attempt to measure competition in the European banking sector directly; instead I rely on contemporary sources and subsequent academic research to support the assumption that the euro was a significant shock to the competitive landscape in the

[11] This concern is not relevant for my first control group which includes EU countries that did not adopt the euro, because they have the same competition policies. Section 3.5 contains the description of control groups.

[12] In my regression analysis, I will try to capture this slow moving process by country and time fixed effects.

European banking sector. Contemporary articles in the business press and research papers coming out of policy making institutions, such as the European Central Bank, clearly show that a fundamental restructuring of the European banking industry due to higher competition was a widely expected consequence of the single currency. Even authors who emphasized remaining tax, regulatory and legal barriers as impediments to a genuine single market acknowledged that the euro would give a boost to cross-border sales of financial products and to cross-border competition. Table 3.2 illustrates these contemporary expectations by quotes from the business press.[13] There is also a recent strand of empirical industrial organization (IO) research which has tried to detect quantitative changes in the degree of the competition in the European banking sector over time. I briefly summarize the challenges of this literature, and I conclude that we have some weak evidence in favor of increased competition after the introduction of the euro.

There are several reasons why establishing a large currency area might foster competition among banks in all member countries. This section focuses on three pro-competition aspects of the creation of the single European currency: increased substitutability of banks' products, more intense competition from the part of capital markets, and more uniform institutional and infrastructural background.

A crucial factor influencing the market power of any company is the degree of substitutability of its product. In the standard monopolistic competition setting this preference parameter is the sole determinant of monopoly power. Each country having its own currency, the national banking sectors enjoyed a certain level of protection from foreign competition because the products banks offer (e.g. loans and deposits) are hard to substitute if they are denominated in different currencies. First, it is harder to compare and evaluate products in different currencies. Second, contracts in foreign currency might involve additional administrative, accounting and transaction costs. Finally, a foreign currency asset or liability exposes the holder to exchange rate risk. The creation of the currency area demolished these natural currency barriers; as a result, each bank had to face a much larger set of potential competitors from abroad. As Table 3.2 reveals, the increased potential competition due to easier substitutability is echoed in many contemporaneous discussions.[14]

In addition to boosting competition within the European banking sector, the creation of the single currency area was also expected to make banks compete with alternative sources of financing, such as the equity and bond markets. Traditionally, corporations in Europe have relied much more heavily on bank loans than their US counterparts. However, the new integrated euro capital market raised the hopes that many companies would be able reach investors at a much lower costs than it was, if at all, possible before. This view about

[13] I systematically searched for relevant articles in the Financial Times, The Economist and The Wall Street Journal between 1 January, 1998 and 31 December, 1999 using the Factiva newspaper database. I applied the following search expression: euro and (bank* near5 competition). The query resulted in 69 articles, and almost all of them contained relevant discussion of the pro-competitive effects of the euro.

[14] Interestingly, not only did the introduction of the euro increase the number of possible rivals within Europe, but the larger market also attracted the attention of US banks. This is evident from an article in the American Banker.

the increasing role of capital markets and the resulting effects on banking is apparent in contemporaneous policy research and in the business press (see Table 3.2, and De Bandt and Davis (2000); BIS (1999); McCauley and White (1997)).

A third feature of the euro that might have promoted more intense competition is the standardization of the institutional and infrastructural background of banks. By institutional background I refer to the available central bank facilities, including the lender of last resort function. National central banks used different operating procedures, which have been replaced by the ECB's instruments for all banks. It is also important to note that, most of the times hidden, but behind every banking system there is a complex machinery that ensures that payments are settled correctly and in a timely manner. The smooth and low cost functioning of clearing houses and payments systems is an essential determinant of the competitiveness of national banking sectors. Before the euro, each country had its own payment infrastructure, and cross-border transactions could be substantially costlier than domestic transactions. With the creation of the Eurozone, these national systems were unified under the oversight of the ECB, which created a level playing field for all banks. Allen et al. (2011) conclude that the harmonization of market infrastructure, such as a uniform cross-border wholesale payment system (TARGET) has certainly contributed to the development of cross-border banking, mainly through lower transaction costs.

The effect of the euro on competition: direct (ex post) evidence

As demonstrated above, there are plenty of reasons to believe that the euro should have raised competitive pressure sharply in the European banking sector, and these beliefs were certainly shared by contemporary observers. Whether competition actually increased under the single currency, is a much harder question to answer. The research on bank competition has evolved mainly in two directions: the structural and non-structural approaches.

The traditional IO theory focuses on market structure, which is reflected in concentration ratios for the largest firms and the Herfindahl index. The literature is based on the assumption that concentration weakens competition by fostering collusive behavior among firms. On the other hand, the non-structural approach posits that factors other than market structure and concentration may affect competitive behavior, such as entry/exit barriers and the general contestability of the market. Non-structural empirical studies do not observe the competitive environment, but they attempt to infer it from behavior. The theoretical advantage of non-structural approaches is that it cannot be assumed *a priori* that concentrated markets are not competitive because contestability may depend on the extent of *potential* competition and not necessarily on market structure. Needless to say, this theoretical advantage comes with an important practical drawback: non-structural results are usually very sensitive to assumptions and estimation methodology.[15]

[15]The most widely employed methodology was developed by Panzar and Rosse (1987). This test is called H-statistics, and is calculated from a reduced form revenue equation by summing up the elasticities of total revenue of the firm with respect to the firm's input prices. Under some assumptions it can be shown that $H <= 0$ for a monopoly, $H = 1$ indicates perfect competition and $0 < H < 1$ indicates monopolistic

Casu and Girardone (2006, 2009) give a comprehensive overview of this literature in the context of European banking, and they also present their own estimates. Generally, structural indicators point to greater concentration after the introduction of the euro. However, the authors argue that it would be misleading to interpret this as a sign of softer competition. On the contrary, the increased concentration has been the consequence of a Europe-wide consolidation process which caused many smaller or less efficient banks disappear through mergers and acquisitions. This wave of M&A's are actually the materialization of the intensified competition brought about by the single market. The findings of non-structural estimations are mixed. For example, Casu and Girardone (2006) find a significantly positive competitive effect of the euro in a sample spanning 1997-2003. However, the same authors' more recent study (Casu and Girardone, 2009), which analyzes competitive pressure between 2000 and 2005, after the introduction of the euro, cannot detect any significant trend. Overall, non-structural results are either insignificant or point to sharper competition in the European banking sector. My conclusion from the empirical IO literature is that it is very hard to measure market competitiveness, and whatever little evidence we have is consistent with the euro bringing about increased competition.

3.5 Data and methodology

My data are taken from the Bankscope database published by Bureau van Dijk. I use annual accounting data for a large sample of individual banks over the time period from 1993 to 2009. Most of my empirical analysis is at the country level, for which I aggregate the micro data up to country-year observations. The variables I use in my analysis are total assets, earning assets, net interest income, and book equity.[16] Bankscope's target audience consists mainly of financial analysts interested in looking at a small sub-sample of banks or countries, not researchers interested in conducting statistical analyses covering many countries and banks included in the dataset. Hence, I had to carefully edit the data before being able to use them for my statistical analysis. Some details of the data and the cleaning method are described in Appendix C.1.

My sample includes all countries that spent at least three years in the eurozone before 2009. On January 1, 1999 the national currencies of eleven European countries ceased to exist, and they were replaced by the new single currency. The initial group was followed by Greece in 2001.[17] In what follows, I will try to detect a euro-effect on equity ratios in these

competition.

[16]Throughout I use the book value of equity and assets to calculate equity ratios. This is different from the Basel regulatory capital requirement, which takes into account the riskiness of assets (risk-weighted assets) and also re-defines equity (e.g. subordinated debt can be included under some circumstances). As a practical matter, only book values are available for a large sample of banks. But even if I had risk adjusted capital ratios, I would be hesitant to trust them more than book values. The recent crisis clearly demonstrated that the risk adjustment banks made are not necessarily more reliable than simple accounting measures.

[17]It is an interesting fact that initially the euro was introduced as non-physical currency, and notes and coins were issued only later. However, the old national currency notes were only used as "strange" units of

twelve countries, which constitutes my treatment group.[18]

Any plausible identification strategy that seeks to capture the effect of the euro requires a comparison group. That is, I need countries that are reasonably similar to the eurozone, but did not introduce the single currency. Ideally, one would consider non-euro countries in the European Union (EU) that share the same competition policies as participants in the euro. There are three EU countries that did not adopt the euro: Denmark, the United Kingdom and Sweden. Members of the European Economic Area (EEA) also have the same EU competition laws, so in principle we could put Norway, Iceland and Lichtenstein into the same category.[19] Unfortunately, sufficient data is unavailable for Iceland and Lichtenstein, so I have to drop them. I also drop Denmark, because although it did not adopt the euro formally, it pegged its national currency to the euro on January 1, 1999.[20] The remaining three countries form my first control group.

Since the ideal control group is quite small, I also consider a second set of countries which might serve as sensible benchmarks. These are industrialized countries that are not part of the EU, but have well-developed banking systems. A third possible control group consists of countries which were not part of the EU at the birth of the euro, but joined in 2004. On *a priori* grounds, this is the least appealing comparison group, because of the big difference in economic development. On the other hand, these countries were required to harmonize their legal systems with EU law by the time of accession, which might have created a somewhat similar competitive environment in banking sector by the early 2000s. For completeness, I will also report some results using this third control group. Table C.1 in Appendix C.1 summarizes the different groups of countries in my sample. Table C.2 in the Appendix C.1 gives summary statistics of my variables.

My empirical methodology to confront the data with the competition hypothesis is closely related to HKS (2011). First, using European data I present some suggestive visual evidence about equity ratios of small versus large banks, and about the relationship between equity ratios and the markup banks charge on loans. Second, I use panel regressions to identify the effect of the euro on banks' capital structure. In particular, I estimate equations of the form

$$y_{ct} = \alpha_c + \delta_t + \beta EURO_{ct} + \gamma X_{ct} + \varepsilon_{ct},$$

where y_{ct} is the country-level equity-to-asset ratio, α and δ are country and time fixed

euros in the interim period.

[18] The next euro adopter was Slovenia in 2007. I do not include it in the sample, because I do not have enough post-euro observations.

[19] EEA countries were unwilling to accept some special EU legislation. Members are allowed to participate in the single market, but in exchange, they are obliged to adopt all EU legislation related to the single market, except laws on agriculture and fisheries. In particular, they must pursue the same competition policy as the EU.

[20] Most of the large political parties in Denmark favored the introduction of the euro, but the public rejected the idea on a referendum by a slight margin. The political elite did not want to give up the perceived gains of the currency area, so they chose to fix the krone to the euro. With the benefit of hindsight, now it looks a like a great solution.

effects, $EURO_{ct}$ is a dummy variable that switches on the beginning of the year when the euro was introduced in the country, and X are control variables.

3.6 Results

Graphs

As a preliminary step, I use graphs to provide some suggestive evidence which is consistent with the competition hypothesis formalized in my model. The data employed in this section comes from countries that adopted the euro (my treatment group), and covers 1993 and 2009.

First, I study equity ratios by bank size. It is far from obvious how to classify banks into different size categories. Experimenting with various definitions, the overall results did not change much, so I followed practical considerations and defined three size groups as follows: In each year and country, I sorted banks according to their total assets. The lowest 40% was labeled as small; the highest 30% was labeled as large; and the middle 30% was labeled as medium sized banks.[21] Figure 3.3 shows that in the European data we observe the same pattern that was documented for the US. Small banks are able to operate with much higher capital ratios.

This pattern is consistent with the competition hypothesis, assuming small banks have more market power due to relationship banking. However, it would be informative to have some more direct proxy for market power. One very crude measure I was able to calculate from my income statement data is the net interest margin (NIM). The NIM is the difference between the interest income generated by banks and the amount of interest paid out to their lenders, relative to the amount of their earning assets. As a first approximation, it can be thought of as the mark-up that banks charge over their funding cost. Figure 3.4 reveals that small banks, indeed, have higher NIMs, and there is a fairly strong association between the NIM and equity ratio.[22]

If one accepts the assumption that the NIM is an imperfect, but still useful measure of market power, it is interesting to note that it started a declining trend in my treatment group just after the introduction of the euro (Figure 3.5).

[21] Specifying absolute thresholds is problematic for two reasons. First, the typical bank size varies a lot by country as Table C.1 shows. A big bank in Austria might be a small player in the Netherlands. Second, over time the threshold should be adjusted, as the size of the overall banking system increases. The present definition also has the advantage that there are no empty categories in any country. Nevertheless, other definitions yielded qualitatively similar results.

[22] I do not want to push this evidence too far, since the NIM reflects a plethora of factors, and market power is only one of them. For example, small banks' loan portfolios might have a different risk profile, which justifies a higher average NIM even in the absence of monopoly rents. See, for example, Demirguc-Kunt, Laeven and Levine (2004). Also, the positive relationship between the NIM and equity ratios in Figure 3.4 Panel B is driven in large part by cross-sectional variation which is stable over time. That's why it is important that the panel regressions in the next section control for time invariant heterogeneity.

Panel regressions

It is tempting to interpret the previous graphs as supporting my hypothesis, but they can be consistent with many alternative explanations as well. In particular, if an omitted variable is correlated with both NIMs and equity ratios in the cross section of banks or countries, then these pictures would easily obtain without any causal relationship. To mitigate somewhat the omitted variable bias, I will use country and time fixed effects in my regression analysis. These dummies will absorb any country-specific heterogeneity that affects the cross-sectional distribution of equity ratios, and will also account for common time trends in the data.

As a first pass on the data, Figure 3.6 plots the average equity ratio in my treatment group and different control groups before and after the introduction of the euro. I use a six year symmetric window, but the results are not very sensitive to this choice. Two observations are evident from the figure. First, in all three control groups the average equity ratio increased, while it remained unchanged in the treatment group. Second, the old EU members who did not adopt the euro and the non-EU developed countries seem to be much more similar, on average, to my treatment group. This confirms my *a priori* expectations about the validity of the various control groups.

Table 3.3 reports the panel regression results. Under the coefficients *p-values* are reported in parentheses. The first three columns confirm the evidence that the graphs conveyed. The euro indicator variable is significant at the 5% level for my first two control groups, and slightly insignificant for the new EU countries.

However, this reduced form regression might capture other effects than what I am after here. The most obvious suspect is a mechanical effect coming from increased average bank size. Figure 3.3 clearly shows that big banks operate at much lower equity ratios than their smaller peers. Although this is consistent with a story based on competition, we cannot rule out other explanations. If the euro induced a consolidation process in the banking sector, we expect the average bank to increase in size. This would lead to a drop in the country-level equity ratio, because big banks hold less equity. The data shows that average bank size, indeed, started to increase after the introduction of the euro (Figure 3.7).

The last three columns of Table 3.3 control for the average size of banking organizations in the sample countries.[23] The results further strengthen my interpretation of the patterns observed in the data. For the two control groups that seemed more appropriate on *a priori* grounds, the euro coefficient change only marginally and remain significant. For the new member states the coefficient loses its significance.

From the panel regressions I conclude that accounting for country specific factors, common time trends, and the mechanical effect of a consolidation process do *not* eliminate the euro effect evident in the graphs. This still does not prove the workings of the competition channel I proposed, although it is certainly consistent with it. In the next section, I address a possible critique of my methodology, and discuss an alternative channel.

[23] Since the dependent variable (equity ratio) is measured in percentage points, it seemed more appropriate to include the average bank size in logarithmic form. However, using the absolute dollar values yield similar conclusions.

3.7 Robustness

I address two specific issues in this section; one is an alternative explanation of my results which does not rely on the degree of competition, and the other is a possible objection to my identification which exploits the timing of the creation of the euro.

First, even if one believes my reduced form regression results, the euro might have effects through other channels than competition.[24] One candidate is increased moral hazard. With an integrated banking sector and a common central bank, banks might have had the perception that they were offered more implicit government guaranties. This could have exacerbated the moral hazard problem, leading to more risky balance sheets. Notice that it is not necessarily the most straightforward interpretation of the effect of the single currency. Another argument would point out that in a much bigger market fewer banks are "too big to fail", so the moral hazard problem can actually decline. Either way, if moral hazard is the main driver of the results, we would expect that the change in equity ratios is concentrated to large, systematically important banks. After all, they are the likely candidates of a government bailout.

Panel A of Table 3.4 reveals that this is exactly the opposite of what happened. For smaller banks the euro effect is much more pronounced than for large banks. In fact, the overall significant reduction in equity ratios is coming entirely from small and medium sized banks defined by my crude classification. The effect for the largest banks is insignificant. If anything, this is more consistent with the competition hypothesis: probably many of the large banks had been international players and had operated under big competitive pressure even before the introduction of the euro. Hence, the euro might have transformed the competitive environment of smaller banks more profoundly.

A legitimate objection against my empirical strategy is that the euro is not an unexpected shock. In a dynamic version of my model the charter value is equal to the discounted value of future rents, so a positive shock to the degree of future competition should reduce the charter value today. Hence, if banks were entirely aware of the effect of the euro on competition, they should have increased their leverage gradually, and I should not find a significant euro effect at the time of the actual introduction. I have three comments about the possible anticipatory bias of my regressions.[25]

First, I do find a significant euro effect which I could not attribute to any other event that happened around that time. After studying the regulatory environment and reading the contemporary business press, I did not find any other event that can be linked to my statistical findings. Second, although the euro had been a goal of the European Union for a long time, and hence the private sector should have been well prepared in theory, my reading of the contemporary sources paint a more nuanced picture. Business press articles from 1998-1999 give the impression that there was a huge turmoil in the banking sector around the creation of the euro. For example, the consolidation process picked up only after

[24] Recall that I already controlled for the mechanical effect of consolidation.

[25] Notice that the target of this critique is not my mechanism (competition hypothesis), but my empirical strategy. So any other explanations that assign a role to the euro suffer from the same problem.

the actual introduction (Figure 3.7), which seems at odds with a perfectly forward looking model.[26]

Third, a counterfactual experiment might shed some light on this question by asking "What results would my methodology produce, if I pretended that the euro was introduced earlier?". If anticipation is a serious problem, we might expect to find many significant effects in this placebo test. Panel B of Table 3.4 presents these counterfactual regressions both for the whole sample and by bank size. None of the results are significant for the whole sample. The only effects that come out significantly are the ones for small and medium banks if we pretend that the euro was introduced one year earlier. Taking into account that the data for a given year contains all income statements filed between July of that year and June of the next year (see Appendix C.1), it is not unreasonable to credit the actual introduction of the euro for these effects as opposed to anticipations.

Overall, acknowledging the serious limitations of my methodology, I suggest that patterns in the data are compatible with my interpretation of the competition hypothesis as sketched out in the model of Section 3.3.

3.8 Conclusion

This paper provides evidence suggesting that increased competition in the financial sector can contribute to the buildup of leverage. Banks face a trade-off when deciding about their capital structure. On the one hand, asymmetric information problems and the presence of deposit insurance and tax shields render debt a more expansive form of financing than equity. On the other hand, more reliance on debt increases the probability of bankruptcy, which results in the loss of the bank's charter value. The degree of competition affects charter values, and hence changes the way banks balance between these two forces.

I illustrate this fundamental trade-off in a simple model where the advantage of debt is motivated by the tax benefit it offers. In the empirical part, I study the evolution of European banks' capital structure around the introduction of the euro which was arguably a large shock to the competitive environment. My panel analysis uncovers statistically and economically significant effects consistent with the competition hypothesis. Banks, in particular smaller banks, decreased their equity ratios after the introduction of the euro. Complementary evidence suggests this effect can be attributed to increased competitive pressure in the sector.

From a policy perspective, my results lend support to capital regulation. Since financial intermediaries do not internalize fully the systemic risk that their financing decisions create for the rest of the economy, capital structure decisions should not be allowed to be driven by competitive forces.

[26] It is also interesting to note that the final decision about the 11 initial countries that would participate in the single currency was made only on May 3, 1998, at the European Council in Brussels.

Table 3.1: Main regulatory measures affecting the EU banking and financial sectors

Year	Regulation
1977	First Banking Directive
1988	Basle Capital Adequacy Regulation
1988	Deregulation of Capital Movements in the European Monetary System
1989	**Second Banking Directive**
1993	Investment Services Directive
1999	Financial Services Action Plan
2000	Directive on e-Money
2001	Directive on the Reorganisation and Winding-Up of Credit Institutions
2001	Regulation on the European Company Statute
2004	New EU Takeover Directive
2005	White Paper on Financial Services Policy
2007	Capital Requirement Directive (Basle II)
2007	Markets in Financial Instrument Directive

Source: Adapted from Carletti and Vives (2009) and Casu and Girardone (2009).

Table 3.2: Competition-enhancing effects of the euro in the contemporary business press

1. Euro increases cross-border competition (substitutability)
Europe's planned single currency has some obvious nasty consequences for banks and insurers. [...] It reduces the number of currencies to trade, and hence the profits from currency dealing. And it exposes many markets to sharper competition. At present, bankers in Italy, Spain and Portugal, for instance, earn fatter margins on corporate loans than bankers in other lands. **Once loans are denominated in euros, Italian and Iberian borrowers will find foreign bankers ready to offer them a better deal.** (Europe's lovesick bankers. 1998, January 10, *The Economist*)
The euro is expected to accelerate European cross-border deals. By creating the foundation of a pan-European market for capital, it exposes markets to stiffer competition. [...] Big fish in small national markets now fret that the euro has turned them into tiddlers surrounded by potential predators, as vulnerable as the hundreds of mid-sized regional banks in America that have been gobbled up by rivals with a national reach. (Getting in the way. 1999, March 13, *The Economist*)
A number of factors will help to level the playing field between domestic and overseas competitors after the introduction of the euro. **In the first place, the elimination of currency risk will make financial products more accessible across borders.** In addition, banks based in, say, Germany, will be able to use the low-cost retail deposits collected through their domestic branch networks to fund loans in another country which may offer more attractive lending margins. (George Graham. [1998, April 30]. Invaders still face barriers, *Financial Times*)
All this spells trouble for Europe's banks. Not only do they risk being cut out of the process of recycling capital from investors to companies; the euro's arrival will also sharpen competition for whatever traditional banking business remains. **Once denominated in the same currency, loan and deposit rates will be directly comparable across euroland.** The market will be unforgiving to those banks, like Italy's, which manage to cover up their inefficiencies only by charging fat interest margins. (Capital project. 1998, March 30, *Financial Times*)
2. Euro increases competition between banks and capital markets
The euro's less obvious consequences are nastier still. If the euro creates a Europe-wide bond market to rival America's for size and liquidity, big companies will be able to raise money for less. **Smaller companies may be able to venture into the bond market for the first time instead of turning to their banks for financing,** putting further pressure on banks' lending margins. For now, banks still dominate lending in Europe. Their share of the assets of all financial firms is 56% in Britain, 73% in France, 77% in Germany and 78% in Spain. In America, where the corporate bond market has largely replaced bank lending, that figure is 22%. European bankers see the American trend coming their way soon, and it has them worried. (Europe's lovesick bankers. 1998, January 10, *The Economist*)
Certainly, the concept of a vast expansion of both equity and bond markets post-Emu is uncontroversial. [...] With the creation of a Euro-land investor base, and the removal of currency barriers across eleven countries, **there will be a move away from banks, even if national regulations slow the process.** Furthermore, pan-European banking competition should intensify, providing cheaper loans. (Simon Davies [1998, January 26]. In search of a virtuous circle, *Financial Times*)
3. Euro attracts US competition
As Europe moves toward a unified currency, the euro, **U.S. banks are poised to grab a sizable share of what they expect to be one of the most attractive and fastest-growing financial markets of the next decade.** [...] As European banks by the hundreds start competing in the same currency, profits will come under pressure, forcing many to merge. This too is seen as favoring the top U.S. banks, with their leaner operations and strong positions in transaction-intensive businesses.(James R. Kraus [1998, July 16]. U.S. Banks Look Like Leaders in The Euro Race, *American Banker*)

Note: Representative excerpts from a systematic search of the financial press between January 1, 1998 and December 31,1999 using the Factiva newspaper database.

Table 3.3: Panel regressions - Impact of the euro on bank equity ratios

	Old EU (1)	Non EU (2)	New EU (3)	Old EU (4)	Non EU (5)	New EU (6)
EURO	-0.75*	-0.67**	-0.69	-0.66*	-0.57*	-0.51
	(0.020)	(0.005)	(0.064)	(0.047)	(0.020)	(0.106)
LOG (AV_SIZE)	-	-	-	-0.25	-0.23	-1.18*
	-	-	-	(0.272)	(0.209)	(0.028)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Note: Based on an annual country-level panel from 1996-2003 assembled from bank income statements. The table shows regressions of equity-to-assets on a dummy for the euro and the logarithm of average bank size using different control groups (defined in Table A1). The EURO dummy switches on beginning in the year when the country adopted the euro. AV_SIZE is the average of total assets among all banks within each country-year. The dependent variable is the asset-weighted average of the equity-to-assets ratio within each country-year. All regressions include a full set of country and year fixed effects. Under the coefficients, p-values are reported in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels.

Table 3.4: Robustness tests (control group: Old EU)

Panel A: Regression results by bank size

	Small (1)	Medium (2)	Big (3)	All banks (4)
EURO	-2.95***	-1.58**	-0.47	-0.66*
	(0.000)	(0.004)	(0.152)	(0.047)
LOG (AV_SIZE)	-1.53***	-1.31***	-0.46*	-0.25
	(0.000)	(0.000)	(0.044)	(0.272)

Panel B: Anticipation tests

	Small (1)	Medium (2)	Big (3)	All banks (4)
EURO(+1)	-2.06**	-1.15*	-0.32	-0.45
	(0.006)	(0.033)	(0.230)	(0.124)
LOG (AV_SIZE)	-1.63***	-1.18***	-0.14	0.00
	(0.000)	(0.000)	(0.477)	(0.994)
EURO(+2)	-1.37	-0.05	0.39	0.37
	(0.090)	(0.928)	(0.143)	(0.214)
LOG (AV_SIZE)	-1.80***	-0.90***	-0.20	-0.04
	(0.000)	(0.000)	(0.333)	(0.858)
EURO(+3)	0.05	0.23	0.27	0.38
	(0.956)	(0.622)	(0.297)	(0.196)
LOG (AV_SIZE)	-1.35**	0.04	-0.11	0.19
	(0.006)	(0.866)	(0.662)	(0.510)

Note: Panel A: AV_SIZE is the average of total assets among all banks within each country-year and bank size category. The dependent variable is the asset-weighted average of the equity-to-assets ratio within each country-year and bank size category. All regressions include a full set of country and year fixed effects. p-values are reported in parentheses. ***, **, * denote significance at 0.001, 0.01, and 0.05 levels. Panel B: Same as Panel A, but the regression "pretends" that the euro was introduced 1,2 or 3 years earlier. For example, the regression with EURO(+2) shifts the whole sample period 2 years backwards to 1994-2001 and assumes counterfactually that the euro was introduced two year earlier.

Figure 3.1: Balance sheet of the model bank

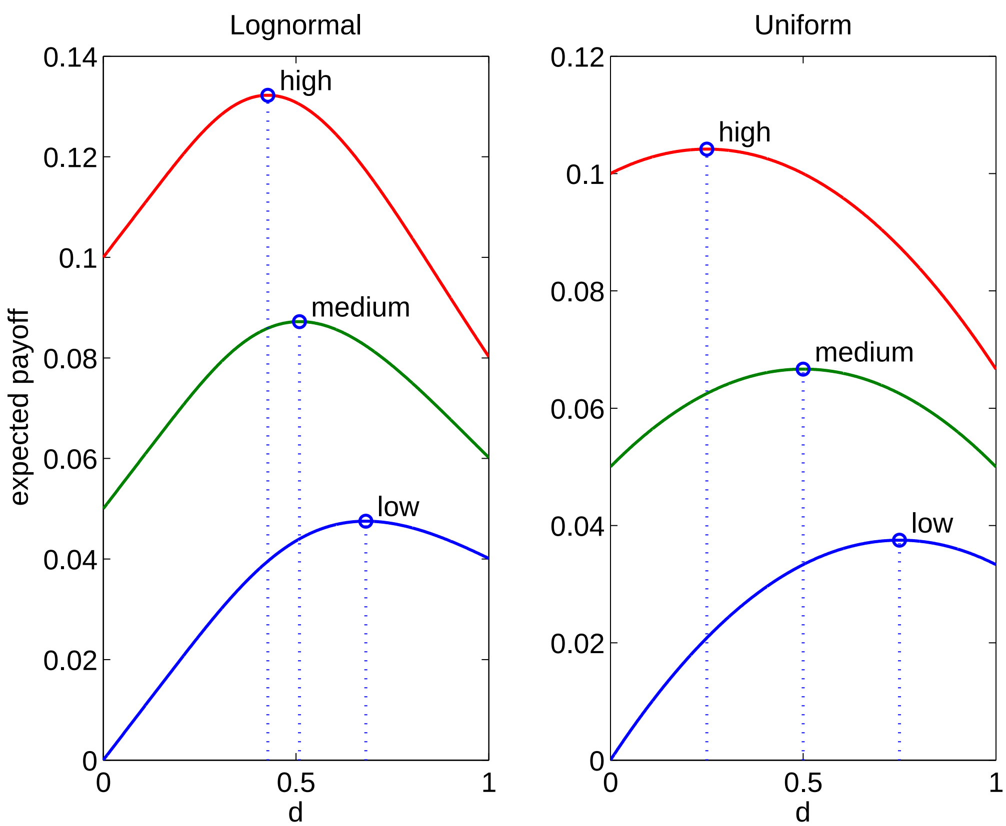
Beginning of period

Assets	Liabilities
A	E_{old} (s shares)
	E_{new} (s' shares)
	Z

End of period

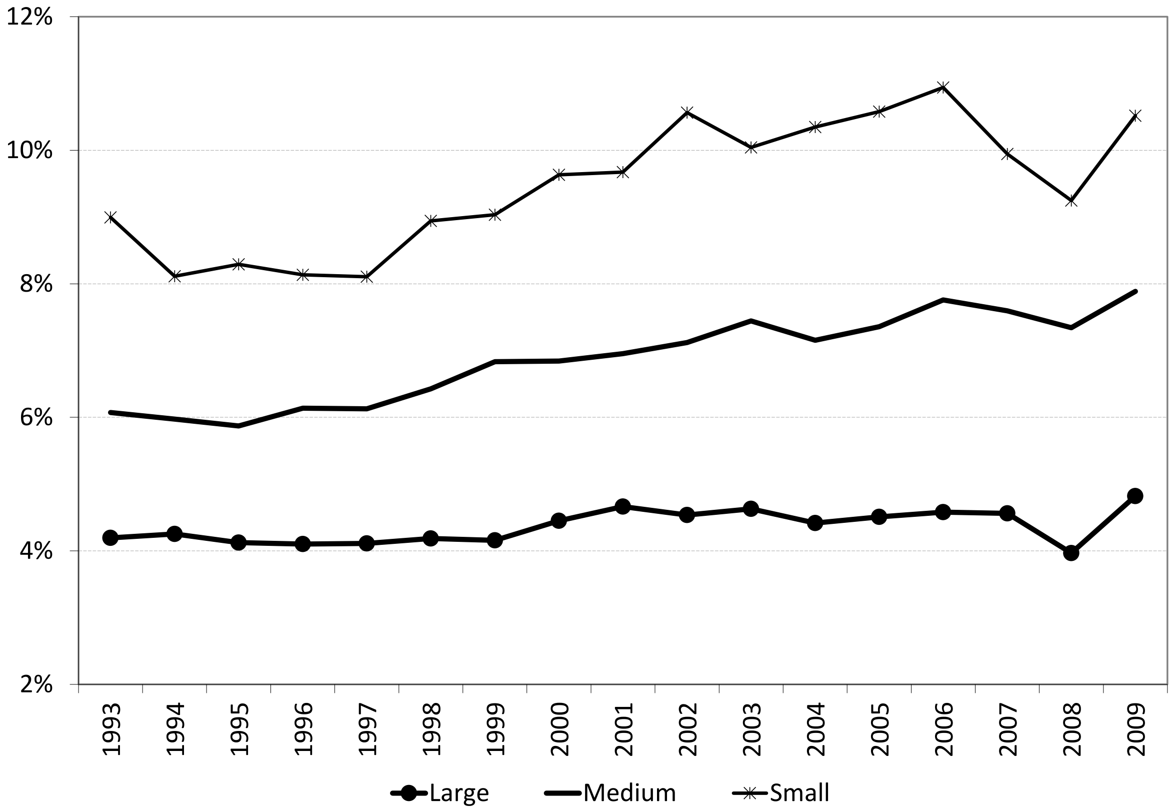
Assets	Liabilities
RA	E ($s + s'$ shares)
	D

Figure 3.2: Charter value and the optimal debt ratio in the model – numerical examples



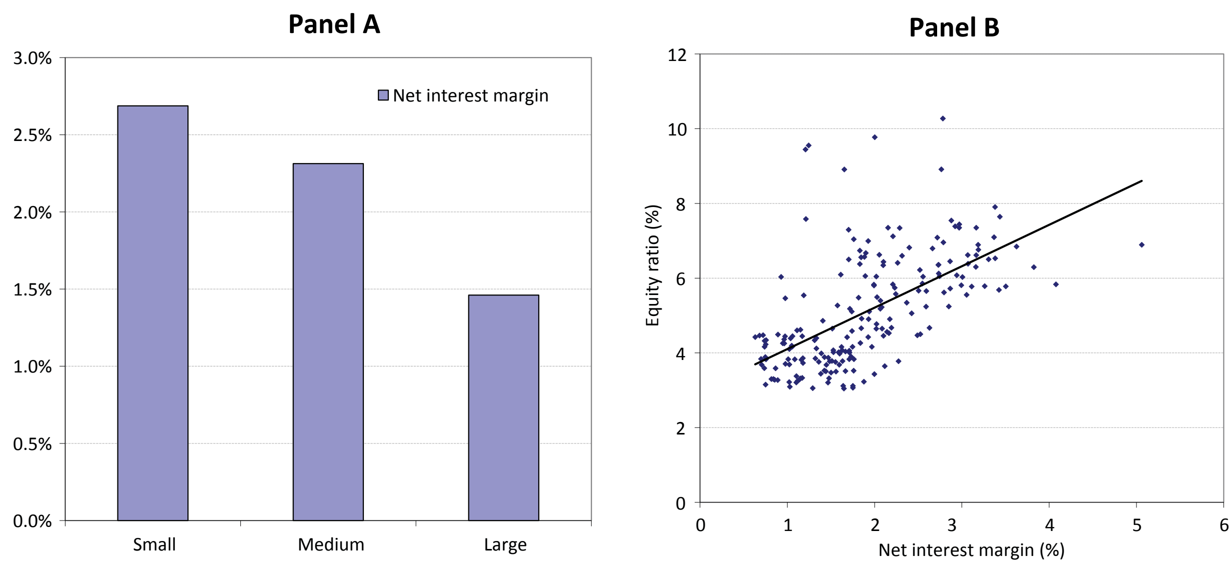
Notes: The figure plots the bank's objective function, equation (XXX), for two standard distributions. Parameters: $\gamma = 0.1, x = 0, 0.25, 0.5$ corresponding to low, medium and high. In the left panel, the return is lognormally distributed: $\log R \sim N(-0.125, 0.5^2)$. In the right panel, the return has a uniform distribution: $R \sim U[0, 1.5]$.

Figure 3.3: Equity ratios by bank size, 1993-2009



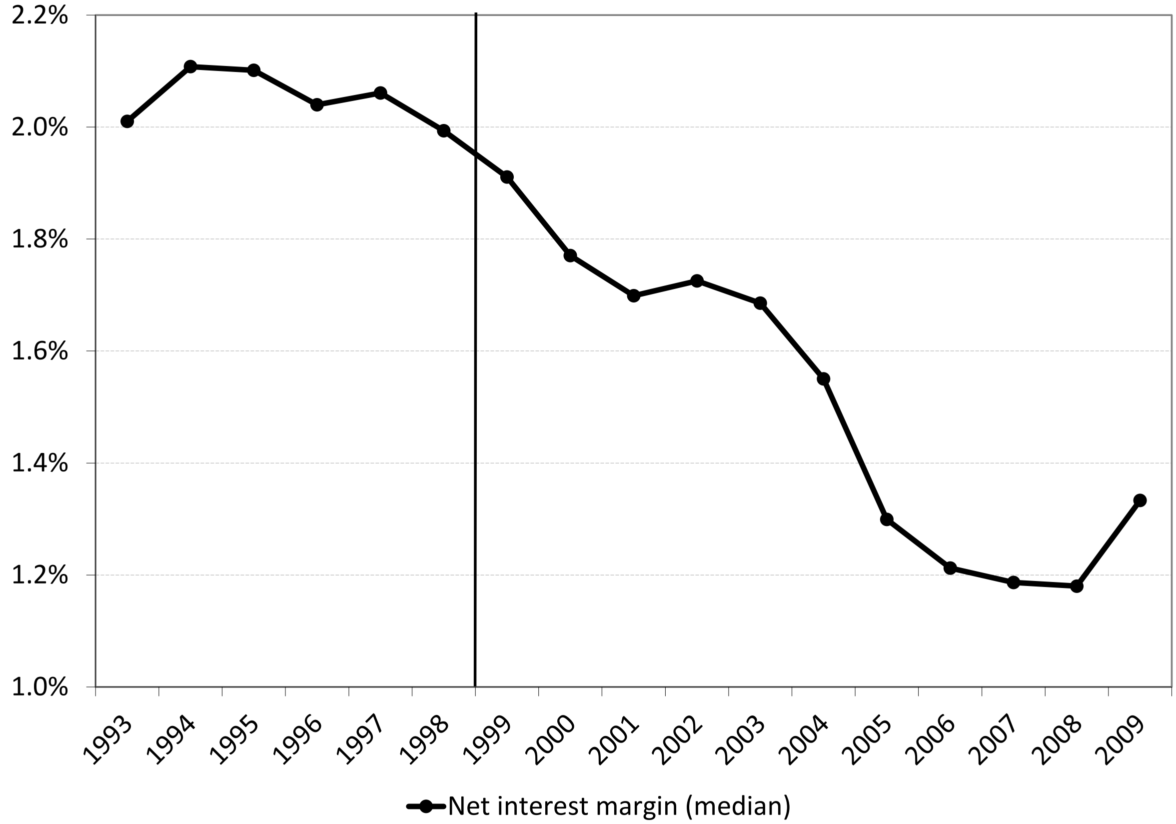
Notes: Aggregate book equity to book assets ratios based on data for all banks in countries adopting the euro (treatment group as defined in Table C.1). Banks are placed into size groups by the following procedure: For a given year and country, banks are ordered according to their total assets. Banks under the 40th percentile are labeled small, between the 40th and 70th percentile medium, and above the 70th percentile large.

Figure 3.4: Net interest margin by bank size (A) and net interest margin vs. equity ratio (B)



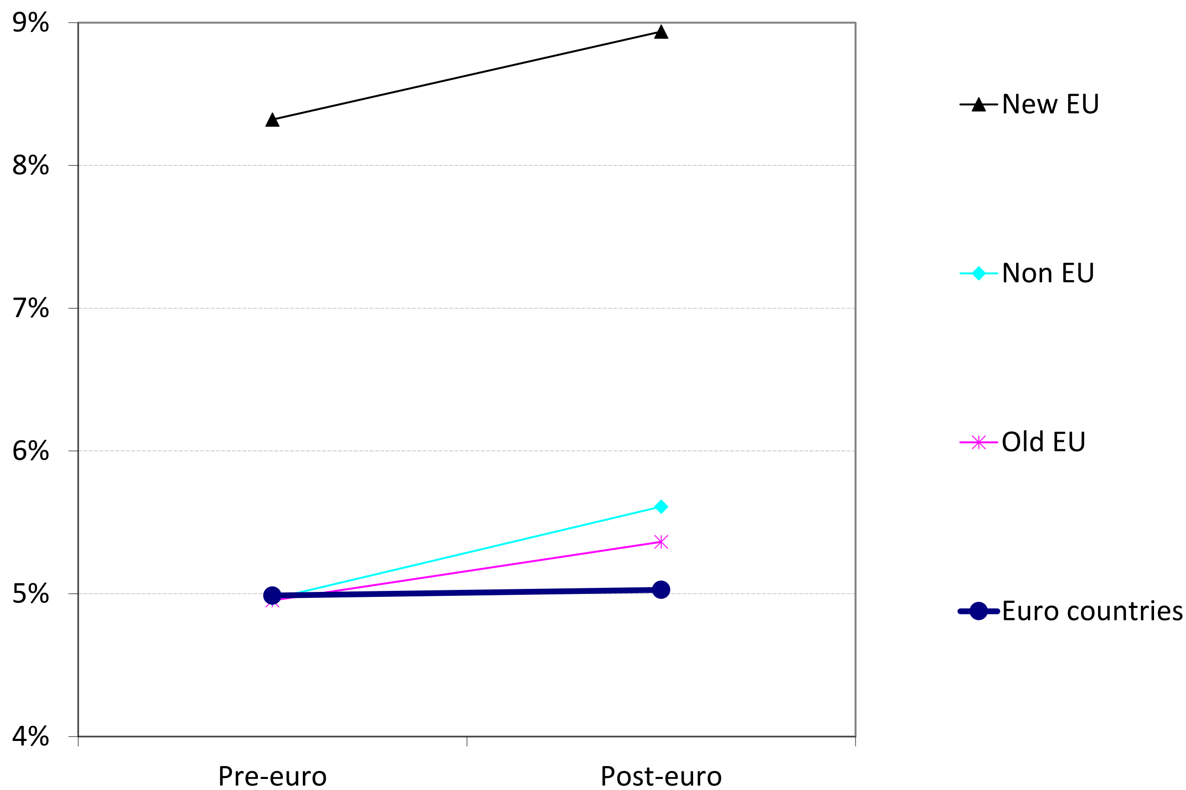
Notes: Panel A uses data for all banks in countries adopting the euro (treatment group as defined in Table C.1). Banks size categories are described under Figure 3.3. Panel B plots country level net interest margins and equity ratios for the treatment group from 1993-2009. Each point corresponds to one country-year observation giving a total of 12*17=204 observations.

Figure 3.5: Evolution of net interest margin, 1993-2009



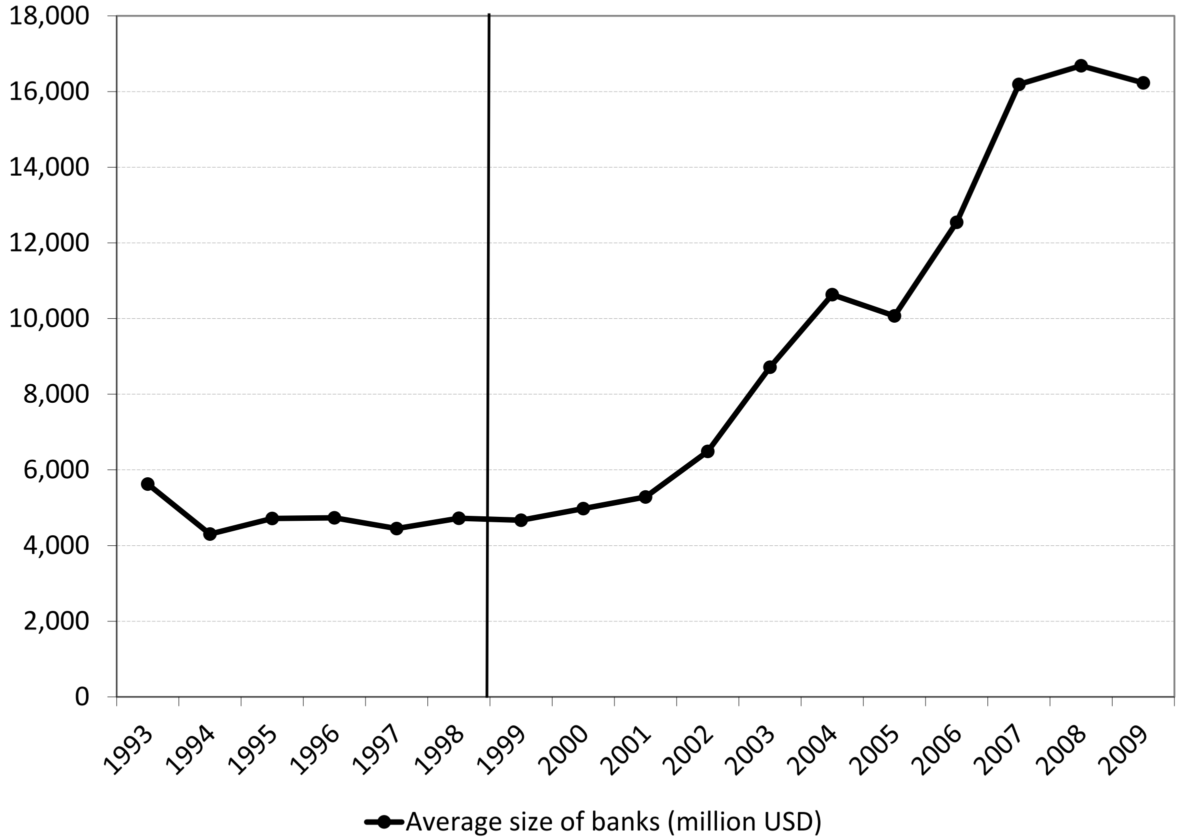
Notes: The figure shows the median of country level net interest margins in the treatment group (defined in Table C.1). The mean is more volatile, but has a similar declining trend. The vertical line corresponds to 1999 when 11 countries adopted the euro.

Figure 3.6: Average equity ratios before and after the euro



Notes: Averages of country-level book equity to book assets ratios. Pre-euro is the average of the three years preceding the euro, and post-euro is the average of the first three years after the introduction of the euro. Country groups are defined in Table C.1.

Figure 3.7: Average bank size measured by assets, 1993-2009



Notes: Average of bank-level book assets based on all banks in the treatment group (defined in Table C.1). The vertical line corresponds to 1999 when 11 countries adopted the euro.

References

Abraham, Katharine G, and Lawrence F Katz. 1986. "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?" *Journal of Political Economy*, 94(3): 507–22.

Admati, Anat R, Peter M DeMarzo, Martin F Hellwig, and Paul C Pfleiderer. 2013. "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity is Not Socially Expensive." Max Planck Institute for Research on Collective Goods Working Paper 2013/23.

Aghion, Philippe, and Gilles Saint-Paul. 1998. "Virtues of Bad Times: Interaction Between Productivity Growth and Economic Fluctuations." *Macroeconomic Dynamics*, 2(3): 322–44.

Alessandria, George, Horag Choi, Joseph P. Kaboski, and Virgiliu Midrigan. 2014. "Microeconomic Uncertainty, International Trade, and Aggregate Fluctuations." National Bureau of Economic Research Working Paper 20616.

Allen, Franklin, Thorsten Beck, Elena Carletti, Philip R. Lane, Dirk Schoenmaker, and Wolf Wagner. 2011. *Cross-border Banking in Europe: Implications for Financial Stability and Macroeconomic Policies.* Centre for Economic Policy Research.

Arellano, Cristina, Yan Bai, and Patrick J. Kehoe. 2012. "Financial Frictions and Fluctuations in Volatility." Federal Reserve Bank of Minneapolis Staff Report 466.

Bachmann, Rüdiger, and Christian Bayer. 2013. "Wait-and-See Business Cycles?" *Journal of Monetary Economics*, 60(6): 704–19.

Bachmann, Rüdiger, and Giuseppe Moscarini. 2012. "Business cycles and endogenous uncertainty." Yale University mimeo.

Bachmann, Rüdiger, Steffen Elstner, and Eric R. Sims. 2013. "Uncertainty and Economic Activity: Evidence from Business Survey Data." *American Economic Journal: Macroeconomics*, 5(2): 217–49.

Basel Committee on Banking Supervision. 2010. "The Basel Committees Response to the Financial Crisis: Report to the G20." http://www.bis.org/publ/bcbs179.pdf.

Bassett, William F., Mary Beth Chosak, John C. Driscoll, and Egon Zakrajsek. 2012. "Changes in Bank Lending Standards and the Macroeconomy." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2012-24.

Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. "Are Technology Improvements Contractionary?" *American Economic Review*, 96(5): 1418–1448.

Baxter, Marianne, and Robert G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series." *The Review of Economics and Statistics*, 81(4): 575–593.

Berger, David, and Joseph Vavra. 2010. "Dynamics of the US price distribution." Yale mimeo.

Bernanke, Ben. 2011. "Implementing a macroprudential approach to supervision and regulation." Speech given at the 47th Annual Conference on Bank Structure and Competition, Chicago, Illinois, May.

BIS. 1999. "BIS Quarterly Review: International Banking and Financial Market Developments." Bank for International Settlements, Basel, Switzerland.

Blinder, Alan S. 2010. "Its Broke, Lets Fix It: Rethinking Financial Regulation." *International Journal of Central Banking*, 6(4): 277–330.

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica*, 77(3): 623–685.

Bloom, Nicholas. 2014. "Fluctuations in Uncertainty." *Journal of Economic Perspectives*, 28(2): 153–76.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry. 2012. "Really Uncertain Business Cycles." National Bureau of Economic Research Working Paper 18245.

Bloom, Nicholas, Paul M. Romer, Stephen J. Terry, and John Van Reenen. 2014. "Trapped Factors and China's Impact on Global Growth." National Bureau of Economic Research Working Paper 19951.

Capistrán, Carlos, and Allan Timmermann. 2009. "Disagreement and Biases in Inflation Expectations." *Journal of Money, Credit and Banking*, 41(2-3): 365–396.

Carletti, Elena, and Xavier Vives. 2009. "Regulation and Competition Policy in the Banking Sector." In *Competition Policy in Europe, Fifty Years of the Treaty of Rome.* , ed. Xavier Vives, 260–283. Oxford:Oxford University Press.

Casu, Barbara, and Claudia Girardone. 2006. "Bank Competition, Concentration And Efficiency in the Single European Market." *Manchester School*, 74(4): 441–468.

Casu, Barbara, and Claudia Girardone. 2009. "Competition Issues in European Banking." *Journal of Financial Regulation and Compliance*, 17(2): 119–133.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. "Chapter 2 Monetary policy shocks: What have we learned and to what end?" In . Vol. 1, Part A of *Handbook of Macroeconomics*, , ed. John B. Taylor and Michael Woodford, 65 – 148. Elsevier.

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy*, 120(1): 116–159.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review*, forthcoming.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia. 2012. "Innocent Bystanders? Monetary Policy and Inequality in the U.S." National Bureau of Economic Research Working Paper 18170.

Cowen, Tyler. 2014. "Automation Alone Isn't Killing Jobs." *The New York Times*.

Croushore, Dean, and Tom Stark. 2001. "A Real-Time Data Set for Macroeconomists." *Journal of Econometrics*, 105(1): 111–130.

Crowe, Christopher W. 2010. "Consensus Forecasts and Inefficient Information Aggregation." International Monetary Fund IMF Working Papers 10/178.

David, Joel M., Hugo A. Hopenhayn, and Venky Venkateswaran. 2014. "Information, Misallocation and Aggregate Productivity." National Bureau of Economic Research Working Paper 20340.

De Bandt, Olivier, and E. Philip Davis. 2000. "Competition, Contestability and Market Structure in European Banking Sectors on the Eve of EMU." *Journal of Banking & Finance*, 24(6): 1045–1066.

Demirguc-Kunt, Asli, Luc Laeven, and Ross Levine. 2004. "Regulations, Market Structure, Institutions, and the Cost of Financial Intermediation." *Journal of Money, Credit and Banking*, 36(3): 593–622.

D'Erasmo, Pablo N, and Hernan J Moscoso-Boedo. 2011. "Intangibles and Endogenous Firm Volatility over the Business Cycle." University of Virginia, Department of Economics Virginia Economics Online Papers 400.

Dovern, Jonas, Ulrich Fritsche, Prakash Loungani, and Natalia Tamirisa. 2015. "Information Rigidities: Comparing Average and Individual Forecasts for a Large International Panel." *International Journal of Forecasting*, 31(1): 144–154.

Driscoll, John C., and Aart C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *The Review of Economics and Statistics*, 80(4): 549–560.

Fama, Eugene F., and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." *The Journal of Finance*, 47(2): 427–465.

Fisher, Jonas D.M., and Ryan Peters. 2010. "Using Stock Returns to Identify Government Spending Shocks." *Economic Journal*, 120(544): 414–436.

Fuster, Andreas, Benjamin Hebert, and David Laibson. 2012. "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing." *NBER Macroeconomics Annual*, 26(1): 1–48.

Gali, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89(1): 249–271.

Gilchrist, Simon, and Egon Zakrajsek. 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review*, 102(4): 1692–1720.

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajsek. 2014. "Uncertainty, Financial Frictions, and Investment Dynamics." National Bureau of Economic Research Working Paper 20038.

Gourinchas, Pierre-Olivier, and Aaron Tornell. 2004. "Exchange Rate Puzzles and Distorted Beliefs." *Journal of International Economics*, 64(2): 303–333.

Gürkaynak, Refet S, Brian Sack, and Eric Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking*, 1(1).

Hall, Robert E. 1991. "Recessions as reorganizations." *NBER macroeconomics annual*, 17–47.

Hamilton, James D. 1996. "This is What Happened to the Oil Price-Macroeconomy Relationship." *Journal of Monetary Economics*, 38(2): 215–220.

Hanson, Samuel G., Anil K. Kashyap, and Jeremy C. Stein. 2011. "A Macroprudential Approach to Financial Regulation." *Journal of Economic Perspectives*, 25(1): 3–28.

Hopenhayn, Hugo A. 1992. "Entry, Exit, and firm Dynamics in Long Run Equilibrium." *Econometrica*, 60(5): pp. 1127–1150.

Ilut, Cosmin, Matthias Kehrig, and Martin Schneider. 2014. "Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News." National Bureau of Economic Research Working Paper 20473.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng. 2013. "Measuring Uncertainty." National Bureau of Economic Research Working Paper 19456.

Kashyap, Anil K, Jeremy C Stein, and Samuel Hanson. 2010. "An analysis of the Impact of Substantially Heightened Capital Requirements on Large Financial Institutions." Mimeo.

Keeley, Michael C. 1990. "Deposit Insurance, Risk, and Market Power in Banking." *The American Economic Review*, 80(5): 1183–1200.

Kilian, Lutz. 2008. "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?" *The Review of Economics and Statistics*, 90(2): 216–240.

Loh, Roger K., and Rene M. Stulz. 2014. "Is Sell-Side Research More Valuable in Bad Times?" National Bureau of Economic Research Working Paper 19778.

Lucas, Robert E. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4(2): 103–124.

Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3): 769–803.

Malmendier, Ulrike, and Stefan Nagel. 2014. "Learning from Inflation Experiences." University of Michigan mimeo.

Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *The Quarterly Journal of Economics*, 117(4): 1295–1328.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. 2004. "Disagreement about Inflation Expectations." *NBER Macroeconomics Annual 2003, Volume 18*, 209–270. The MIT Press.

McCauley, Robert N., and William R. White. 1997. "The Euro and European financial markets." Bank for International Settlements BIS Working Papers 41.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): pp. 1695–1725.

Mertens, Karel, and Morten O. Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review*, 103(4): 1212–47.

Miles, David, Jing Yang, and Gilberto Marcheggiano. 2013. "Optimal Bank Capital." *The Economic Journal*, 123(567): 1–37.

Montiel Olea, José Luis, James Stock, and Mark W. Watson. 2012. "Inference in Structural VARs with External Instruments." Harvard University mimeo.

Morris, Stephen, and Hyun Song Shin. 2002. "Social Value of Public Information." *American Economic Review*, 92(5): 1521–1534.

Moscarini, Giuseppe, and Lones Smith. 2002. "The law of large demand for information." *Econometrica*, 70(6): 2351–2366.

Myers, Stewart C., and Nicholas S. Majluf. 1984. "Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have." *Journal of Financial Economics*, 13(2): 187–221.

Panzar, John C., and James N. Rosse. 1987. "Testing For "Monopoly" Equilibrium." *The Journal of Industrial Economics*, 35(4): pp. 443–456.

Pesaran, M. Hashem, and Martin Weale. 2006. "Survey Expectations." In . Vol. 1 of *Handbook of Economic Forecasting*, , ed. G. Elliott, C.W.J. Granger and A. Timmermann, 715–776. Elsevier.

Phelps, Edmund S. 1968. "Money-Wage Dynamics and Labor-Market Equilibrium." *Journal of Political Economy*, 76(4): 678–711.

Price, David A. 2013. "Interview: Mark Gertler." *Econ Focus*, 11(4): 32–36.

Radner, Roy, and Joseph Stiglitz. 1984. "A Nonconcavity in the Value of Information." *Bayesian models in economic theory*, 5: 33–52.

Rajan, Raghuram G. 1992. "Insiders and Outsiders: The Choice between Informed and Arm's-Length Debt." *The Journal of Finance*, 47(4): pp. 1367–1400.

Ramey, Valerie A. 2011. "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics*, 126(1): 1–50.

Reis, Ricardo. 2006. "Inattentive Producers." *The Review of Economic Studies*, 73(3): 793–821.

Romer, Christina D., and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94(4): 1055–1084.

Romer, Christina D., and David H. Romer. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review*, 100(3): 763–801.

Sims, Christopher A. 1998. "Stickiness." *Carnegie-Rochester Conference Series on Public Policy*, 49(0): 317–356.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3): 665–690.

Smets, Frank, and Raf Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, 1(5): 1123–1175.

Stock, James H., and Mark W. Watson. 2008. "What's New in Econometrics Time Series, Lecture 7: Structural VARs." NBER Summer Institute Minicourse 2008.

Stock, James H., and Mark W. Watson. 2012. "Disentangling the Channels of the 2007-09 Recession." *Brookings Papers on Economic Activity*, 44(1 (Spring): 81–156.

U.S. Treasury. 2009. "Principles for Reforming the U.S. and International Regulatory Capital Framework for Banking Firms." U.S. Treasury white paper. http://www.ustreas.gov/press/releases/docs/capital-statement 090309.pdf.

Woodford, Michael. 2001. "Imperfect Common Knowledge and the Effects of Monetary Policy." National Bureau of Economic Research Working Paper 8673.

Appendix A

Appendix to Chapter 1

A.1 Robustness checks

Construction of dispersion measures

The empirical results reported in the paper are reasonably robust to changes in the construction of the dispersion measures. I experimented with different cutoff values for inclusion in the sample (5, 10, 15 and 25 years). I also examined the effect of removing the predictable common component from the firm-level variables before calculating their cross-sectional dispersion. In this exercise, I used the Fama and French (1992) three-factor model to capture the common variation in stock returns.[1] For sales growth numbers, I regressed firm-level variables on GDP growth and the change in unemployment rate, and defined the residuals from these regressions as the idiosyncratic component of sales growth. Reporting all the results would be impossible, but we can get an idea of robustness by looking at the co-movement of these alternative dispersion series. Table A.1 and A.2 show the correlation matrix of the resulting dispersion measures. The correlation coefficients are very high, usually above 0.9. This is the reason why none of my empirical results changes substantially if I construct my dispersion measures differently.

Variation in the set of external instruments

This section documents that the baseline results from the structural VAR estimation in 1.2 are robust to changes in the set of instruments. I focus on the impulse response function of dispersion measures to output level-shocks.

First, I take advantage of the GMM methodology and test for over-identifying restrictions within the baseline specifications. The Hansen J-tests cannot reject the null that the five external instruments identify the same causal relation (p=0.36 for sales growth and p=0.41 for stock return dispersion). Next, I carry out a sensitivity analysis by re-estimating the

[1] I downloaded the factors from Kenneth French's website.

models with dropping one instrument at a time. In the main text, I have shown that excluding the Gilchrist and Zakrajsek (2012) excess bond premium has little impact on the estimated impulse responses. Figure A.1 demonstrates that excluding any other instruments does not alter the results either. The estimated responses are remarkably close to the baseline model. The overidentification tests and the evidence in Figure A.1 suggest that the different instruments identify the same causal relationship, and that my results are not driven by one particular shock series.

Second, I re-estimate the SVAR models with a completely different set of instruments. As explained in the main text, data limitations forced me to include only one shock from each category in the baseline specifications. I chose the shock series with the longest available time series to ensure sufficient overlap and variation for estimation. Here I present the results with a completely different set of instruments: Kilian (2008) oil price shock; Gürkaynak, Sack and Swanson (2005) monetary policy shock; Gali (1999) productivity shock; Romer and Romer (2010) fiscal shock; Bassett et al. (2012) credit supply shock. Figure A.2 compares the resulting impulse response functions with the baseline specification. Using the alternative instruments, negative level-shocks still trigger an immediate rise in dispersion, although the effect is smaller and much less precisely estimated (error bands are not shown on the graph). The weaker results are not surprising, since some of these shocks have very short time series (see Table 1.1). The qualitative results, however, are completely consistent with the baseline estimations.

A.2 Identification with External Instruments

The derivation here follows closely Montiel Olea, Stock and Watson (2012).

Assume that Y_t is a column vector of r variables, which follows a VAR

$$A(L)\,Y_t = \eta_t,$$

where η_t is the vector of VAR innovations. The r innovations in η_t are linear combinations of r structural shocks ε_t, so that

$$\eta_t = H\varepsilon_t = [H_1 \;\cdots\; H_r]\begin{pmatrix}\varepsilon_{1t}\\ \vdots \\ \varepsilon_{rt}\end{pmatrix}, \tag{A.1}$$

where H_1 is the first column of H, ε_{1t} is the first structural shock, and so forth. The structural shocks are assumed to be uncorrelated, so $\Sigma_{\varepsilon\varepsilon} = E(\varepsilon_t\varepsilon_t')$ is diagonal. We also assume, as is standard in the structural VAR literature, that the system described in equation (A.1) is invertible, so that the structural shocks can be expressed as linear combinations of the innovations:

$$\varepsilon_t = H^{-1}\eta_t.$$

Our ultimate goal is to derive the impulse response function of Y_t with respect to the structural shocks, which is equivalent to identifying the columns of the H matrix. Here we consider the problem of identifying the effect of a single shock, which for convenience we take to be the first shock ε_{1t}. Suppose we have a vector of k instrumental variables Z_t, not included in Y_t, such that

$$\text{(i) } E\left(\varepsilon_{1t}Z_t'\right) = \alpha' \neq 0 \text{ (relevance)}$$
$$\text{(ii) } E\left(\varepsilon_{jt}Z_t'\right) = 0, \quad j = 2, ..., r \text{ (exogeneity).}$$

Condition (i) says that Z_t is correlated with the shock of interest, ε_{1t}; that is, Z_t is a relevant instrument. Condition (ii) says that Z_t is uncorrelated with the other structural shocks. By conditions (i) and (ii), Z_t is correlated with η_t only because it is correlated with ε_{1t}.

Conditions (i) and (ii) imply that

$$E\left(\eta_t Z_t'\right) = E\left(H\varepsilon_t Z_t'\right) = [H_1 \ \cdots \ H_r]\begin{pmatrix} E\left(\varepsilon_{1t}Z_t'\right) \\ \vdots \\ E\left(\varepsilon_{rt}Z_t'\right) \end{pmatrix} = H_1\alpha',$$

where the first equality follows from equation (A.1) and the final equality from conditions (i) and (ii). The instrument Z_t thus identifies H_1 in the population up to scale and sign.

The scale and sign of H_1 are set by normalizing the shock to have a unit impact on a given variable; for example, an oil price shock is normalized so that a 1-unit positive shock increases the (log) oil price by 1 unit. If we order the variable associated with ε_{1t} first, then our normalization implies that a unit positive increase in ε_{1t} should have the impact effect of a unit positive increase in η_{1t}, and thus in Y_{1t}. Formally,

$$\text{(iii) } H_{11} = 1 \text{ (unit shock normalization),}$$

where H_{11} is the first element of H_1. After imposing (iii), we obtain

$$E\left(\eta_t Z_t'\right) = \begin{pmatrix} 1 \\ H_{1\cdot} \end{pmatrix}\alpha', \tag{A.2}$$

so that α' is identified by $E\left(\eta_{1t}Z_t'\right)$ and the remaining $r-1$ rows of (A.2) identify $H_{1\cdot}$.

The GMM of $H_{1\cdot}$ obtains by solving the sample analogs of the moment condition (A.2). Using the Kronecker-product notation, the GMM objective function can be written as

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left((\hat{\eta}_t \otimes Z_t) - (\begin{bmatrix} 1 \\ H_{1\cdot} \end{bmatrix} \otimes \alpha)\right)' W^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left((\hat{\eta}_t \otimes Z_t) - (\begin{bmatrix} 1 \\ H_{1\cdot} \end{bmatrix} \otimes \alpha)\right)$$

where W is a weight matrix and $\hat{\eta}_t$ is the vector of estimated residuals from the VAR. Since the VAR coefficients are estimated consistently, replacing η_t with the estimated innovations

$\hat{\eta}_t$ does not change the asymptotic properties of the GMM estimator.[2] Once we have an estimate of H_1, it is straightforward to construct the IRF of Y_t after an exogenous shock to ε_{1t}.

In my empirical work, I implement this identification scheme in a bivariate VAR with

$$Y_t = [\log GDP_t, d_t]',$$

where d_t is one of the cross-sectional dispersion measures. I use the aggregate shock series as external instruments (Z_t) for the structural innovation in the level of GDP.

A.3 Learn to avoid disasters (Illustrative model)

The model in the main text generates countercyclical information acquisition through time-variation in the *cost* of information. However, the *benefit* from more precise information is also likely to be larger in bad times if it helps avoid disastrous outcomes. Suppose, for example, that resorting to external financing is much more expensive than using internal cash flows. When aggregate output is low, the probability of hitting the external financing threshold is high. Thus, any increase in expected profit is especially beneficial because the firm can save on the external financing cost too. We can think of other distortions, such as collateral constraints, bankruptcy costs or a minimum efficient scale of production, that can induce strong nonlinearities at low levels of activity. In a booming economy these constraints are slack, and suboptimal decisions are unlikely to move a firm close to them. In contrast, recessions are periods when even the consequences of small mistakes loom large because they can push firms against these constraints. This reasoning suggests that information is more valuable in bad times.

Here I develop this intuition in a static partial equilibrium setting where external financing is costly.

The firm's problem

Assume that a continuum of firms set their prices. Firm i's (real) profit from normal operation is given by $y - (p_i - z_i)^2$, where p_i is the chosen relative price, z_i is some idiosyncratic shock and y is an aggregate state that shifts the profit function.[3] The cross-sectional distribution of z_i is $N(0, \alpha^{-1})$, which is also the prior belief of the firm. There are two frictions in the market. First, information about z_i is imperfect, and it comes in the form of a noisy signal $s = z_i + \beta^{-0.5}\varepsilon$ with $\varepsilon \sim N(0, 1)$. Second, the firm incurs an extra cost of $f > 0$ if its normal operating profits go below zero. One can think of f as the fixed cost of external financing that the firm has to pay in order to meet financial obligations such as wage

[2] For technical details, see Olea, Stock and Watson (2012).

[3] This can be thought of as a log-quadratic approximation of the standard monopolistic competition model.

payments.[4] Hence, total profit is

$$\pi_i = y - (p_i - z_i)^2 - I_{[y-(p_i-z_i)^2<0]} f,$$

where I is the indicator function.

After observing the signal, the firm wants to set a price to maximize expected profits

$$\max_{p_i} E_s[\pi_i] = y - E_s[(p_i - z_i)^2] - P_s[y - (p_i - z_i)^2 < 0] f. \tag{A.3}$$

The s subscript signifies that the decision is conditional on the signal. The two terms reflect the two goals of the firm. It wants to set the price that maximizes the expected value of normal operating profit, but also wants to avoid crossing the zero-profit threshold. In general, there can be a trade-off between these two goals. However, with a symmetric payoff function and symmetric distribution this trade-off will be absent. The optimal price will be given by the firm's posterior belief about z_i:

$$p_i = \mu(s) \equiv \frac{\beta s}{\alpha + \beta},$$

where $\mu(s) = E[z_i|s]$.[5]

The interesting question is how much information the firm will acquire to guide its price-setting decision, and how information-gathering incentives depend on the economic environment. Assume that obtaining the signal costs $w\beta$, where w is the real wage. Similarly to the model in the main text, this captures the idea that gathering more information or processing existing information better requires more time. Plugging in $\mu(s)$ into (A.3), and taking expectations, we get the firm's ex ante expected profit before the realization of the signal:

$$\pi^e(\beta) = y - (\alpha + \beta)^{-1} - 2\Phi\left(-\sqrt{y(\alpha + \beta)}\right) f,$$

where Φ denotes the standard normal cdf.

Expected profit obviously increases in the degree of information acquisition. To find the optimal signal precision, the firm balances the cost and benefit of information gathering. This is illustrated in Figure A.3. The solid horizontal line represent the marginal cost of information, which is given by w. The solid downward sloping marginal benefit of information (MBI) curve represents the increase in expected profit after acquiring an extra unit of signal precision.[6]

[4]Besides one-time fees related to borrowing, a fixed cost can also capture the idea that seeking external finance exacerbates the conflict of interest between managers and outside financiers. See, for example, Eisfeldt and Muir (2014) and the references therein for empirical evidence on the important role of fixed external financing costs.

[5]Formally, the first order condition is given by

$$-2(p_i - \mu(s)) - (\alpha + \beta)^{0.5}\left[\phi\left(\frac{p_i - \mu(s) - \sqrt{y}}{(\alpha + \beta)^{-0.5}}\right) - \phi\left(\frac{\mu(s) - p_i - \sqrt{y}}{(\alpha + \beta)^{-0.5}}\right)\right] f = 0$$

where ϕ is standard normal pdf. It is easy to see that the only solution is $p_i = \mu(s)$ and that this satisfies the second order condition.

[6]The marginal benefit of information (MBI) is given by

Information acquisition, dispersion and aggregate variables

The cross-sectional dispersion of prices is directly linked to how much firms attend to the idiosyncratic shock. When attention is low, they mostly make decisions based on the aggregate state of the economy, so their choices will be similar. The more they take into account firm-level conditions, the more dispersed the resulting optimal prices will be. Formally, the relationship between attention and cross-sectional dispersion is given by

$$Var_i\left(p_i\right) = \frac{\beta}{\alpha+\beta}\alpha^{-1}.$$

In this static model, we can capture "bad times" by low w and y.[7] Figure A.3 depicts the comparative statics with respect to these parameters. The downward shift of the marginal cost curve increases information acquisition. The intuition is the same as in the model of the main text: information is cheaper, so the firm buys more of it. However, this model highlights a new mechanism, which shifts the MBI curve out, and increases optimal information acquisition even further. Figure A.4 elaborates on the intuition for this outward shift by comparing the gains from information acquisition in good and bad times. The bottom part of the figure shows realized operating income as a function of the unknown fundamental assuming that the firm sets its price optimally given its information. The top section displays the firm's subjective uncertainty about z for two levels of information acquisition (low and high). An important benefit of better information is that the probability of hitting the zero-profit threshold and incurring the penalty can be reduced. The decrease in this probability is depicted by the shaded region between the two density functions. Evidently, in a boom the very bad outcome is already quite unlikely, so the additional gain is small. In contrast, when aggregate conditions are bad, reducing the probability of hitting this constraint can bring about substantial gains.

This simple model reiterates the main messages from the text and somewhat extends the intuition for the mechanism. *First*, in a world of imperfect information, there are reasons to expect cyclical variation in the optimal level of information acquisition. Attending to information requires time and mental capacity which otherwise could be used as input into other productive activities. The opportunity cost of these resources is likely to be lower in bad times. In addition, bad times carry the possibility of highly undesired outcomes with strong nonlinearities in payoffs, so deviating from the optimal action can be more costly than in good times. This suggests that the benefit from better information is also

---

$$\frac{d\pi^e(\beta)}{d\beta} = (\alpha+\beta)^{-2} + \sqrt{y}\,(\alpha+\beta)^{-0.5}\,\phi\left(\sqrt{y}\,(\alpha+\beta)^{0.5}\right) f$$

which is strictly positive and decreasing in β.

[7] In the data, real wages are procyclical, although moderately. The Smets and Wouters (2003) New Keynesian model implies that output and real wages move in the same direction after many structural shocks (e.g. productivity, monetary policy). According to their estimated parameters, these shocks play an important enough role in business cycles that the simulated data generated by the model also features a positive correlation.

larger in bad times. The movement of costs and benefits over the business cycle affects the incentives to acquire information which has not been studied in the rational inattention literature. *Second*, the level of attention to idiosyncratic shocks has a direct impact on the dispersion of realized outcomes. If idiosyncratic conditions are overlooked, then decision-makers' information sets are more correlated, which leads to more correlation in actions. The combination of these forces introduces a new channel which is able to generate the well-documented countercyclicality of various cross-sectional dispersion measures.

A.4 Detailed derivations

Profit maximization

As I show in the main text, the optimal price is given by the usual mark-up over expected marginal cost formula:

$$\begin{aligned} p(s) &= E_s\left[\frac{w}{\rho A z}\right] = \frac{w}{\rho A E_s\left[z^{-1}\right]^{-1}} \\ &= \frac{w}{\rho A \bar{z}\left(s\right)}. \end{aligned}$$

To calculate $\bar{z}\left(s\right)$, we use the standard Bayesian updating formula for normal variables. Recall that the prior distribution of z is $\log z \sim N(0.5\alpha^{-1}, \alpha^{-1})$, and the signal is given by $s = \log z + \beta^{-0.5}\varepsilon$ with $\varepsilon \sim N(0,1)$. Hence, the posterior of $\log z$ is also normal, and we can write

$$\begin{aligned} \bar{z}(s;\beta) &= E[z^{-1}|s]^{-1} = E[e^{-\log z}|s]^{-1} = [e^{-E[\log z|s]+0.5Var[\log z|s]}]^{-1} \\ &= \exp\left(\frac{\alpha\left(0.5\alpha^{-1}\right)+\beta s}{\alpha+\beta} - 0.5(\alpha+\beta)^{-1}\right) \\ &= \exp\left[\frac{\beta}{\alpha+\beta}s\right]. \end{aligned}$$

Plugging in the optimal price into the profit function, we can calculate the value of a firm with signal s:

$$\begin{aligned} v\left(s\right) &= E_s\left[\left(p\left(s\right)-\frac{w}{Az}\right)\left(\frac{p\left(s\right)}{P}\right)^{-\sigma}Y\right] = \left(p\left(s\right)-\frac{w}{A\bar{z}\left(s\right)}\right)\left(\frac{p\left(s\right)}{P}\right)^{-\sigma}Y \\ &= \left(1-\rho\right)p\left(s\right)\left(\frac{p\left(s\right)}{P}\right)^{-\sigma}Y = \frac{1}{\sigma}\left(\frac{p\left(s\right)}{P}\right)^{1-\sigma}R \\ &= \frac{R}{\sigma}\left(\rho A\bar{z}(s)\frac{P}{w}\right)^{\sigma-1}. \end{aligned}$$

Taking ex ante expectations, we derive the firm's expected profit before observing the signal:

$$\pi^e(\beta) = E\left[v\left(s\right)\right] = \frac{R}{\sigma}(\rho A\tilde{z}(\beta)\frac{P}{w})^{\sigma-1},$$

where we use that $\bar{z}\left(s\right)$ is itself a lognormal random variable, so that

$$\begin{aligned}
\tilde{z}(\beta) &= E\left[\bar{z}(s;\beta)^{\sigma-1}\right]^{\frac{1}{\sigma-1}} = E\left[e^{(\sigma-1)\log\bar{z}(s)}\right]^{\frac{1}{\sigma-1}} = \left[e^{(\sigma-1)E[\log\bar{z}(s)]+0.5(\sigma-1)^2 Var[\log\bar{z}(s)]}\right]^{\frac{1}{\sigma-1}} \\
&= e^{E[\log\bar{z}(s)]+0.5(\sigma-1)Var[\log\bar{z}(s)]} = \exp\left[\frac{\beta}{\alpha+\beta}0.5\alpha^{-1} + 0.5\left(\sigma-1\right)\left(\frac{\beta}{\alpha+\beta}\right)^2\left(\alpha^{-1}+\beta^{-1}\right)\right] \\
&= \exp\left[0.5\frac{\beta}{\alpha+\beta}\left(\alpha^{-1}+\left(\sigma-1\right)\frac{\beta}{\alpha+\beta}\frac{\alpha+\beta}{\alpha\beta}\right)\right] \\
&= \exp\left[\frac{1}{2}\frac{\beta}{\alpha+\beta}\sigma\alpha^{-1}\right].
\end{aligned}$$

Proof of Proposition 1. First calculate $d\tilde{z}(\beta)/d\beta$:

$$\frac{d}{d\beta}\tilde{z}(\beta) = \tilde{z}(\beta)\left[\frac{1}{2}\sigma\alpha^{-1}\frac{(\alpha+\beta)-\beta}{(\alpha+\beta)^2}\right] = \frac{1}{2}\sigma\tilde{z}(\beta)(\alpha+\beta)^{-2}$$

Using this result, the first derivative of the expected profit function is

$$\begin{aligned}
\frac{d}{d\beta}\pi^e(\beta) &= \frac{R}{\sigma}\left(\rho A\frac{P}{w}\right)^{\sigma-1}\left(\sigma-1\right)\tilde{z}(\beta)^{\sigma-2}\frac{d\tilde{z}(\beta)}{d\beta} \\
&= \frac{1}{2}\left(\sigma-1\right)R\left(\rho A\tilde{z}(\beta)\frac{P}{w}\right)^{\sigma-1}(\alpha+\beta)^{-2} > 0 \qquad \text{(A.4)}
\end{aligned}$$

and the second derivative is

$$\begin{aligned}
\frac{d^2}{d\beta^2}\pi^e\left(\beta\right) &= \frac{1}{2}(\sigma-1)R(\rho A\frac{P}{w})^{\sigma-1}\left[(\sigma-1)\tilde{z}(\beta)^{\sigma-2}\frac{d\tilde{z}(\beta)}{d\beta}(\alpha+\beta)^{-2} - 2\tilde{z}(\beta)^{\sigma-1}(\alpha+\beta)^{-3}\right] \\
&= \frac{1}{2}(\sigma-1)R\left(\rho A\tilde{z}\left(\beta\right)\frac{P}{w}\right)^{\sigma-1}(\alpha+\beta)^{-4}\left[\frac{1}{2}(\sigma-1)\sigma - 2(\alpha+\beta)\right]
\end{aligned}$$

which is smaller than zero if and only if $\frac{1}{4}(\sigma-1)\sigma-\alpha<\beta$. Hence, if $\frac{1}{4}(\sigma-1)\sigma-\alpha<0$, then the expected profit function is increasing and strictly concave on $[0,\infty)$. Given the assumed properties of $G(\beta)$, this implies that the firm has a unique optimal degree of information acquisition, and the optimum can be fully described by the first order condition. ■

Aggregation and equilibrium

I impose $w=1$. Assume that all firms choose the same signal precision β in equilibrium. The i.i.d. assumption about the idiosyncratic shock and the signal noise implies that the firms' subjective prior distribution over the signal coincides with the distribution of realized signal values across firms. In particular, the realized signal values will be distributed as $s \sim N(0.5\alpha^{-1}, \alpha^{-1}+\beta^{-1})$. Let us denote the corresponding pdf by $f(s)$. Then the aggregate price level can be written as

$$
\begin{aligned}
P &= \left[\int_0^1 p_i^{1-\sigma} di\right]^{\frac{1}{1-\sigma}} = \left[\int_{-\infty}^{\infty} p(s)^{1-\sigma} f(s)\, ds\right]^{\frac{1}{1-\sigma}} = \left[\int_{-\infty}^{\infty} \left(\frac{1}{\rho A \bar{z}(s;\beta)}\right)^{1-\sigma} f(s)\, ds\right]^{\frac{1}{1-\sigma}} \\
&= \frac{1}{\rho A}\left[\int_{-\infty}^{\infty} \bar{z}(s;\beta)^{\sigma-1} f(s)\, ds\right]^{\frac{1}{1-\sigma}} = \frac{1}{\rho A} E\left[\bar{z}(s;\beta)^{\sigma-1}\right]^{\frac{1}{1-\sigma}} \\
&= \frac{1}{\rho A \tilde{z}(\beta)},
\end{aligned}
$$

where I used the definition of $\tilde{z}(\beta)$.

Combining this expression with the production technology (1.4), the demand function (1.3), and the price-setting rule (1.5), we can derive the amount of labor used in the production of goods. Let $f(s,z)$ denote the joint distribution of the signal and the productivity shock (both across firms and from the perspective of an individual firm before observing the signal). Then we can write:

$$
\begin{aligned}
L_{goods} &= \int_0^1 l_i di = \int_0^1 \frac{\left(\frac{p_i}{P}\right)^{-\sigma} Y}{A z_i} di \\
&= \frac{Y}{A}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(\frac{\tilde{z}(\beta)}{\bar{z}(s;\beta)}\right)^{-\sigma} z^{-1} f(s,z)\, dz ds = \frac{Y}{A}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} z^{-1} f(z|s)\, dz\right]\left(\frac{\tilde{z}(\beta)}{\bar{z}(s;\beta)}\right)^{-\sigma} f(s)\, ds \\
&= \frac{Y}{A}\int_{-\infty}^{\infty} E[z^{-1}|s]\left(\frac{\tilde{z}(\beta)}{\bar{z}(s;\beta)}\right)^{-\sigma} f(s)\, ds = \frac{Y}{A} E\left[\bar{z}(s;\beta)^{-1}\left(\frac{\tilde{z}(\beta)}{\bar{z}(s;\beta)}\right)^{-\sigma}\right] \\
&= \frac{Y}{A}\tilde{z}(\beta)^{-\sigma} E\left[\bar{z}(s;\beta)^{\sigma-1}\right] = \frac{Y}{A}\tilde{z}(\beta)^{-\sigma}\tilde{z}(\beta)^{\sigma-1} \\
&= \frac{Y}{A\tilde{z}(\beta)}.
\end{aligned}
$$

Total labor demand is the sum of labor used for goods production and for information acquisition:

$$
L^D = L_{goods} + L_{info} = \frac{Y}{A\tilde{z}(\beta)} + G(\beta).
$$

After substituting for the aggregate price level, the household's labor supply condition is

$$
L^S = \chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}} \left(\rho A \tilde{z}(\beta)\right)^{\frac{1}{\gamma-1}}.
$$

Labor supply approaches ∞ as $Y \to 0$ and it is monotonically decreasing to 0 as $Y \to \infty$. Labor demand is monotonically increasing Y. Hence, there is a unique intersection which determines equilibrium output for any common information acquisition strategy β. All the other aggregate variables follow easily as a function of β.

Now assume that in equilibrium all firms choose β^{EQ}. An individual firm (whose measure is zero) will set its signal precision β^* to satisfy the first order condition in (1.10):

$$\frac{d}{d\beta}\pi^e\left(\beta^*;\beta^{EQ}\right) = G'\left(\beta^*\right),$$

where the inclusion of β^{EQ} in π^e indicates that the firm's expected profit depends on other firms' information acquisition strategy through their impact on aggregate variables. Substituting aggregate variables into (A.4) and imposing the fixed point property of Nash-equilibrium ($\beta^* = \beta^{EQ}$), we obtain

$$\frac{1}{2}(\sigma-1)R\left(\beta^{EQ}\right)\left(\alpha+\beta^{EQ}\right)^{-2} = G'\left(\beta^{EQ}\right), \tag{A.5}$$

where again I made it explicit that aggregate expenditures, R, depend on β^{EQ} in equilibrium.

Proof of Proposition 2. We want to show that there is a unique β^{EQ} that satisfies the above condition. Given the assumptions on the G function, it suffices to show that (i) $R(0) > 0$ and (ii) $R(\beta)$ is decreasing in β. (i) $R(0) = P(0)Y(0) > 0$ follows immediately from our earlier derivations. (ii) We will use the implicit function theorem repeatedly to show that $dR/d\beta < 0$. Start from the following equilibrium relationships:

$$P = \frac{1}{\rho A\tilde{z}(\beta)} \tag{A.6}$$

$$0 = \frac{Y}{A\tilde{z}(\beta)} + G(\beta) - \chi^{\frac{-1}{\gamma-1}}Y^{-\frac{\theta}{\gamma-1}}P^{\frac{1}{1-\gamma}}. \tag{A.7}$$

First note that the $G(\beta)$ term will always decrease the value of $dY/d\beta$. That is, for any increase in β, the increase in Y would be bigger if we did not have to sacrifice some labor to acquire more information. In order to bound this derivative from above, I will omit this term and use the notation $\overline{dY/d\beta}$ to indicate that this is an upper bound. If we ignore the $G(\beta)$ term, then β only affects aggregate equilibrium variables through $\tilde{z}(\beta)$. Hence, we can write

$$\frac{dR}{d\beta} = \frac{dP}{d\beta}Y + P\frac{dY}{d\beta} < \frac{dP}{d\beta}Y + P\overline{\frac{dY}{d\beta}} = \left(\frac{dP}{d\tilde{z}}Y + P\overline{\frac{dY}{d\tilde{z}}}\right)\frac{d\tilde{z}}{d\beta},$$

where $d\tilde{z}/d\beta > 0$, so it is enough to establish that the term in parenthesis is nonnegative. From (A.6) we can directly calculate $dP/d\tilde{z} = -P/\tilde{z} < 0$, which implies

$$\frac{dP}{d\tilde{z}}Y + P\overline{\frac{dY}{d\tilde{z}}} = \frac{dP}{d\tilde{z}}\left(Y - \tilde{z}\overline{\frac{dY}{d\tilde{z}}}\right).$$

■

Now let us calculate $\overline{dY/d\tilde{z}}$ by applying the implicit function theorem to (A.7):

$$\frac{\overline{dY}}{d\tilde{z}} = \frac{Y}{\tilde{z}} \frac{\frac{1}{A\tilde{z}} + \frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}{\frac{1}{A\tilde{z}} + \theta\frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}.$$

Using this expression, we can write

$$Y - \tilde{z}\frac{\overline{dY}}{d\tilde{z}} = Y\left[1 - \underbrace{\frac{\frac{1}{A\tilde{z}} + \frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}{\frac{1}{A\tilde{z}} + \theta\frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}}_{<1}\right] > 0,$$

where I used the assumption $\theta > 1$. Putting everything together, we have shown that $dR/d\beta < 0$. Note that the $\theta > 1$ condition is certainly sufficient to establish the desired monotonicity property, but it is absolutely not necessary. In fact, the upper bound on the derivative is always slack because I ignored the $G(\beta)$ term.

Information acquisition, dispersion and the aggregate state

In what follows, aggregate variables R, P and Y should be understood as their equilibrium values satisfying (A.6), (A.7) and (A.5) for a given set of parameters. The goal is to do comparative statics with respect A which represents the aggregate state of the economy.

Proof of Proposition 3. From (A.5), we can write

$$\frac{d\beta^{EQ}}{dA} = -\frac{\frac{1}{2}(\sigma-1)\frac{dR}{dA}\left(\alpha+\beta^{EQ}\right)^{-2}}{\frac{1}{2}(\sigma-1)\left[\frac{dR}{d\beta^{EQ}}\left(\alpha+\beta^{EQ}\right)^{-2} - 2R\left(\alpha+\beta^{EQ}\right)^{-3}\right] - G''\left(\beta^{EQ}\right)}.$$

Since we already showed in Proposition 2 that $dR/d\beta < 0$ and $G''\left(\beta^{EQ}\right) \geq 0$ by assumption, the sign of the above expression only depends on the sign of dR/dA. Observing that $dP/dA = -P/A < 0$, we can write

$$\begin{aligned}\frac{dR}{dA} &= \frac{dP}{dA}Y + P\frac{dY}{dA} = \frac{dP}{dA}\left(Y - A\frac{dY}{dA}\right)\\ &= \frac{dP}{dA}Y\left(1 - \underbrace{\frac{\frac{1}{A\tilde{z}} + \frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}{\frac{1}{A\tilde{z}} + \theta\frac{1}{\gamma-1}\chi^{\frac{-1}{\gamma-1}} Y^{-\frac{\theta}{\gamma-1}-1} P^{\frac{1}{1-\gamma}}}}_{\leq 1}\right) \leq 0.\end{aligned}$$

This implies that $d\beta^{EQ}/dA < 0$. ■

Proof of Proposition 4. Using firms' optimal price from (1.5), the log price change between $t-1$ and t for any firm i can be written as

$$\Delta \log p_{it} = \log \frac{p_{it}(s_{it};\beta_t)}{p_{i,t-1}(s_{i,t-1};\beta_{t-1})} = \log \frac{A_{t-1}\bar{z}(s_{i,t-1};\beta_{t-1})}{A_t \bar{z}(s_{it};\beta_t)}$$

Substituting $\bar{z}\,(s;\beta)$ from (1.6), we obtain

$$\Delta \log p_{it} = \frac{\beta_{t-1}}{\alpha+\beta_{t-1}} s_{i,t-1} - \frac{\beta_t}{\alpha+\beta_t} s_{i,t} + \log \frac{A_{t-1}}{A_t}$$

Since the s_{it} is i.i.d. across firms, the cross sectional variance of the above expression is

$$\begin{aligned} Var_i\left(\log \frac{p_{it}}{p_{i,t-1}}\right) &= \left(\frac{\beta_{t-1}}{\alpha+\beta_{t-1}}\right)^2 Var_i(s_{i,t-1}) + \left(\frac{\beta_t}{\alpha+\beta_t}\right)^2 Var_i(s_{it}) \\ &= \left(\frac{\beta_{t-1}}{\alpha+\beta_{t-1}}\right)^2 \frac{\alpha+\beta_{t-1}}{\alpha\beta_{t-1}} + \left(\frac{\beta_t}{\alpha+\beta_t}\right)^2 \frac{\alpha+\beta_t}{\alpha\beta_t} \\ &= \left[\frac{\beta_{t-1}}{\alpha+\beta_{t-1}} + \frac{\beta_t}{\alpha+\beta_t}\right]\alpha^{-1}. \end{aligned}$$

■

Table A.1: Robustness of dispersion measures (sales growth)

		Raw data				Idiosyncratic comp.			
	years	5	10	15	25	5	10	15	25
Raw data	5	1							
	10	0.99	1						
	15	0.98	0.99	1					
	25	0.87	0.89	0.92	1				
Idiosyncratic component	5	0.93	0.92	0.89	0.81	1			
	10	0.94	0.93	0.91	0.82	0.99	1		
	15	0.95	0.94	0.93	0.85	0.98	0.99	1	
	25	0.88	0.89	0.9	0.91	0.86	0.88	0.91	1

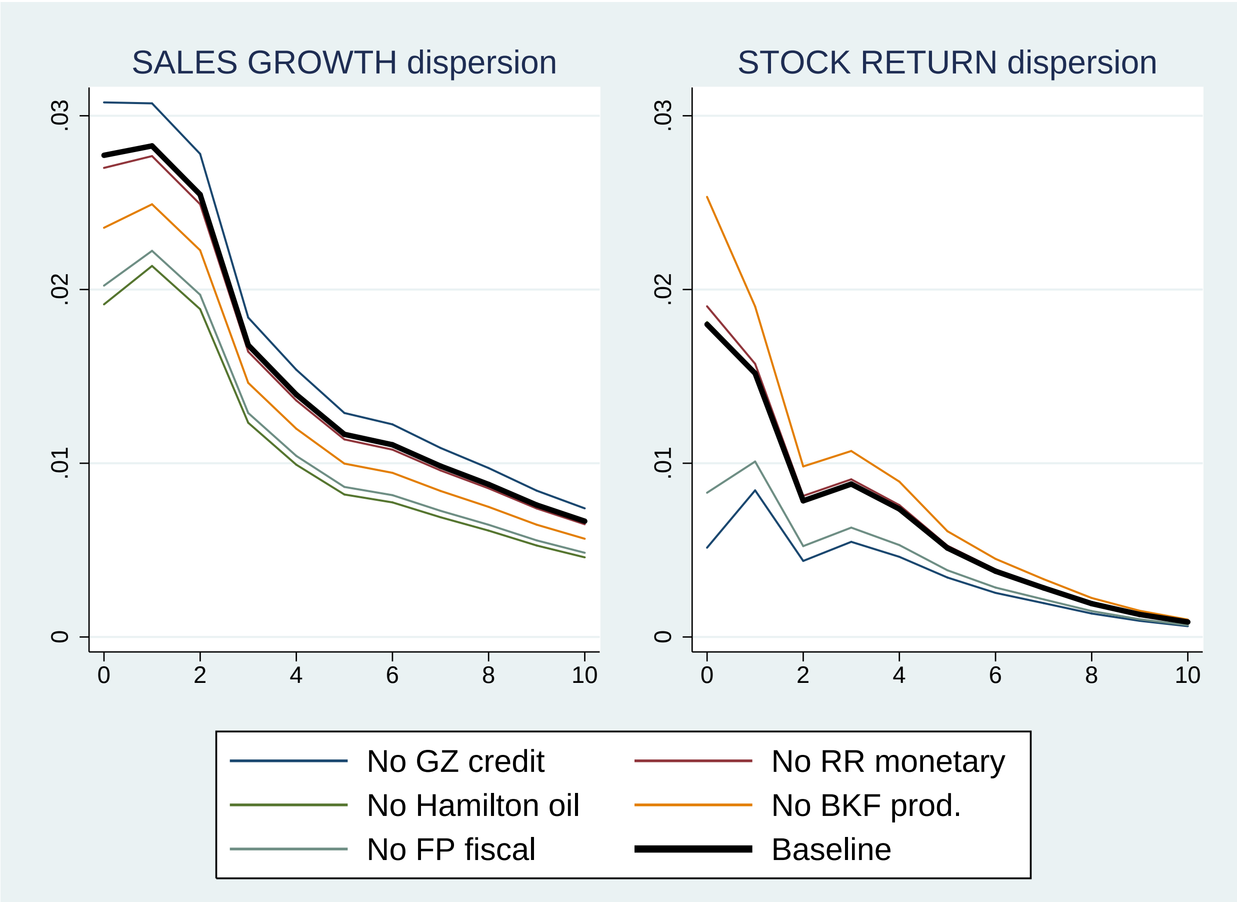
Notes: The table reports the correlation of sales growth dispersion measures constructed in slightly different ways. The rows labeled “Raw data” use unfiltered firm-level sales growth numbers. The rows labeled “Idiosyncratic component” filter out the common variation before calculating cross-sectional dispersion. I regress firm-level sales growth on GDP growth and the change in unemployment rate, and define the idiosyncratic component as the residual. The “years” column indicates the cut-off value of firm tenure for inclusion in the sample.

Table A.2: Robustness of dispersion measures (stock returns)

		Raw data				Idiosyncratic comp.			
	years	5	10	15	25	5	10	15	25
Raw data	5	1							
	10	0.99	1						
	15	0.98	0.99	1					
	25	0.93	0.96	0.98	1				
Idiosyncratic component	5	0.91	0.89	0.86	0.81	1			
	10	0.93	0.92	0.91	0.87	0.98	1		
	15	0.92	0.92	0.92	0.9	0.96	0.99	1	
	25	0.89	0.9	0.92	0.92	0.9	0.95	0.98	1

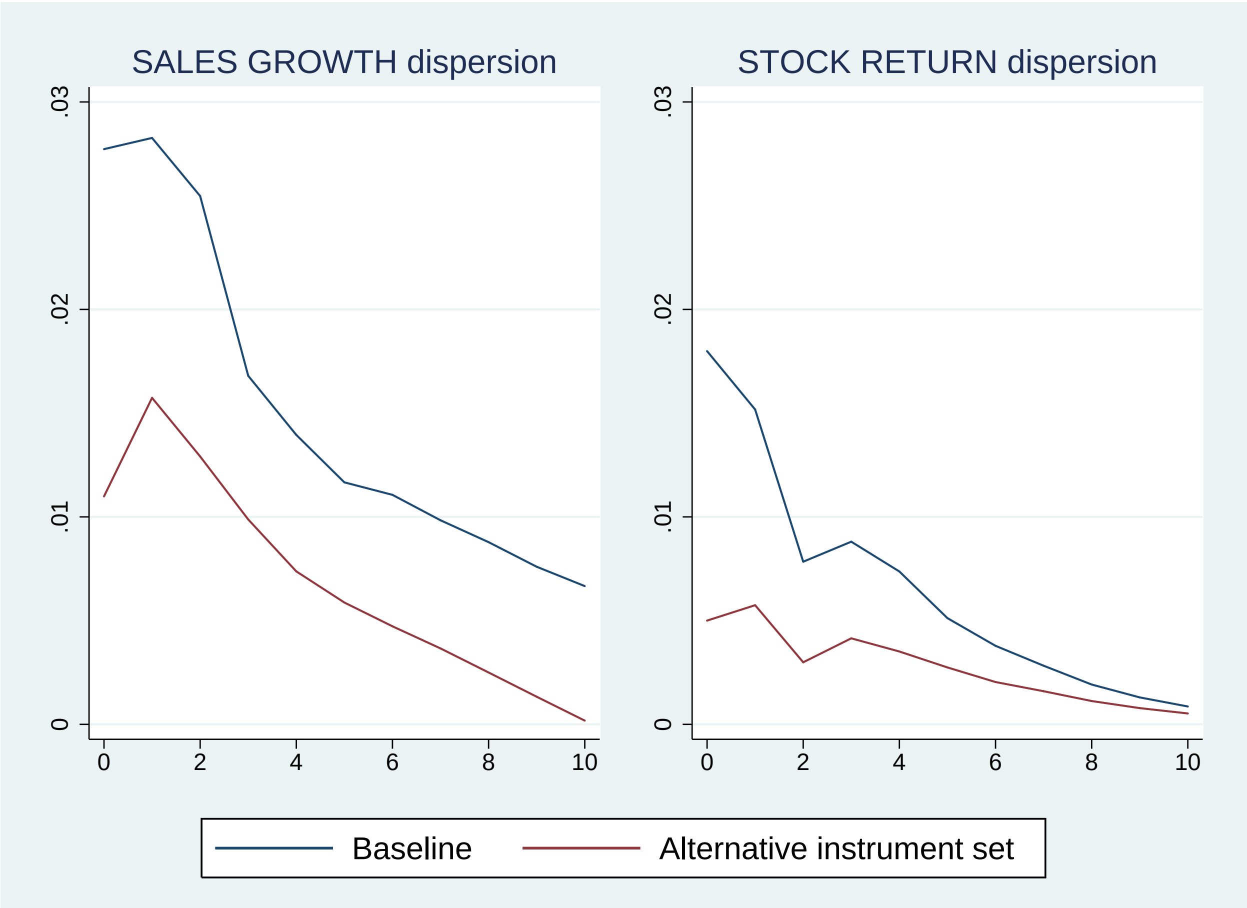
Notes: The table reports the correlation of stock return dispersion measures constructed in slightly different ways. The rows labeled "Raw data" use unfiltered firm-level stock return numbers. The rows labeled "Idiosyncratic component" filter out the common variation before calculating cross-sectional dispersion. I regress firm-level stock returns on the Fama and French (1992) factors, and define the idiosyncratic component as the residual. The "years" column indicates the cut-off value of firm tenure for inclusion in the sample.

Figure A.1: Robustness: Dropping one external instrument at a time



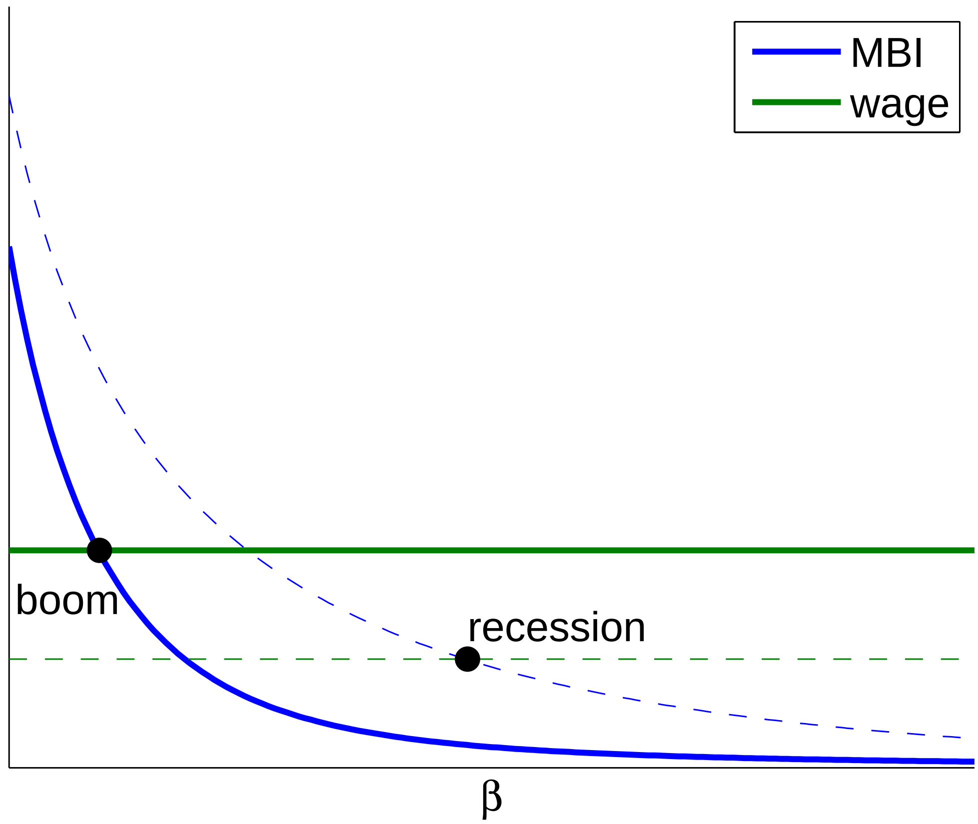
Notes: The thick black lines show the response of sales growth dispersion (left) and stock return dispersion (right) to an output level-shock in the baseline SVARs. (See section 1.2 and the notes below Figure 1.11 for the baseline specifications.) Each colored thin line depicts the same response in a slightly modified SVAR where one external instrument is excluded from the estimation. The legend identifies the excluded shock: Gilchrist and Zakrajsek (2012) credit supply shock; Hamilton (1996) oil price shock; Fisher and Peters (2010) fiscal shock; Romer and Romer (2004) monetary policy shock; Basu, Fernald and Kimball (2006) productivity shock.

Figure A.2: Robustness: Alternative instrument set



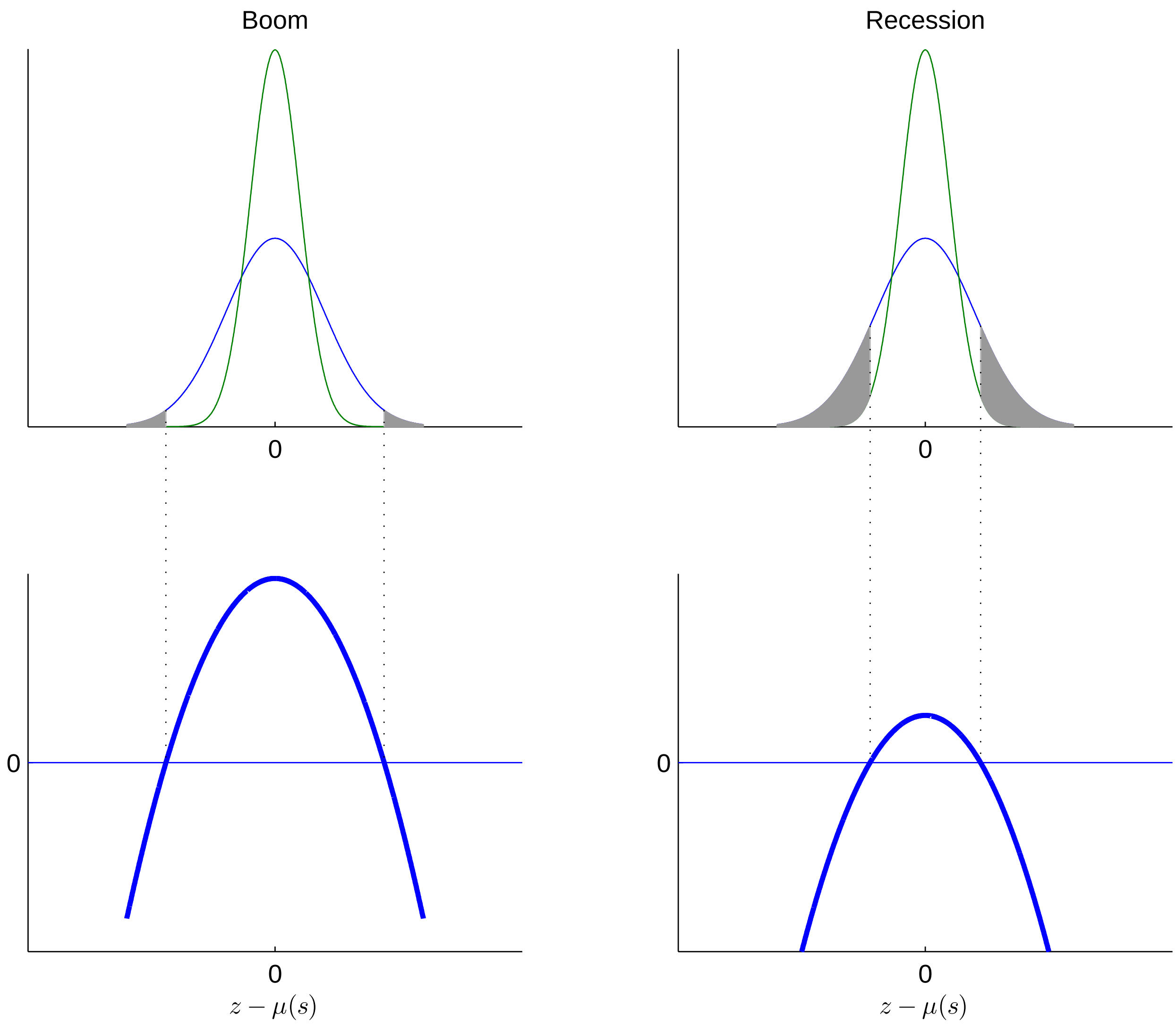
Notes: The blue lines show the response of sales growth dispersion (left) and stock return dispersion (right) to an output level-shock in the baseline SVARs. (See section 1.2 and the notes below Figure 1.11 for the baseline specifications.) The red lines depict the same response estimated with a completely different set of instruments: Kilian (2008) oil price shock; Gürkaynak, Sack and Swanson (2005) monetary policy shock; Gali (1999) productivity shock; Romer and Romer (2010) fiscal shock; Bassett et al. (2012) credit supply shock. The qualitative results are very similar to the preferred baseline specification.

Figure A.3: Optimal information acquisition with costly external financing



Notes: Optimal signal precision in the partial equilibrium model of Appendix A.3. The solid lines show the marginal benefit of information (blue) and the cost of labor used in information acquisition (green). The optimum is determined by the intersection. The dashed lines show the effect of a "recession", which is captured with lower y and lower w. The benefit of information increases in bad times (MBI shifts out), because better information reduces the probability that firms have to resort to costly external financing. In booms, this probability is very low for any level of information acquisition (see Figure A.4).

Figure A.4: Benefit from more precise information



Notes: The gain from information acquisition is countercyclical in the partial equilibrium model of Appendix A.3. The graph illustrates the benefit from better information for two levels of y: high ("boom") and low ("recession"). The bottom part shows operating income as a function of the unknown fundamental. The top section displays the firm's subjective uncertainty about z for two levels of information acquisition (low and high). The shaded region gives the decrease in the probability of hitting the zero-profit threshold. In a boom the bad outcome is already unlikely, so the additional gain is small. In a recession, however, better information can generate substantial gains.

Appendix B

Appendix to Chapter 2

This technical appendix contains the derivations that were omitted in the main text.

B.1 Forecast errors and forecast revisions in the model

Here I derive the structural relationships that provide the basis for interpreting the reduced form regressions (Equations (2.1), (2.6) and (2.10) in the text.)

Combining the heuristic updating rule in equation (2.3) and the signal process in (2.2), simple algebra yields the following relationship

$$z_t = \frac{1}{G} z_{it|t} - \frac{1-G}{G} z_{it|t-1} - \omega_{it}.$$

Using this expression and the AR(1) process for z_t, we can derive the relationship between ex-post forecast errors and ex-ante forecast revisions at the individual level:

$$\begin{aligned}
z_{t+h} - z_{i,t+h|t} &= \rho^h \left(z_t - z_{it|t}\right) + \sum_{j=1}^{h} \rho^{h-j} v_{t+j} \\
&= \rho^h \left[\frac{1-G}{G}\left(z_{it|t} - z_{it|t-1}\right) - \omega_{it}\right] + v_{t+h,t} \\
&= \frac{1-G}{G}\left(z_{i,t+h|t} - z_{i,t+h|t-1}\right) - \rho^h \omega_{it} + v_{t+h,t}.
\end{aligned}$$

Averaging across agents, we obtain the analogous relationship for the consensus forecast:

$$z_{t+h} - \overline{z_{t+h|t}} = \frac{1-G}{G}\left(\overline{z_{t+h|t}} - \overline{z_{t+h|t-1}}\right) + v_{t+h,t}. \tag{B.1}$$

Similar derivations yield the relationship between current and past forecast revisions:

$$\begin{aligned}
z_{i,t+h|t} - z_{i,t+h|t-1} &= \rho^h \left(z_{it|t} - \rho z_{i,t-1|t-1}\right) \\
&= \rho^h \left[G\left(z_t + \omega_{it}\right) + (1-G)\, z_{it|t-1} - \rho\left(G\left(z_{t-1} + \omega_{it-1}\right) + (1-G)\, z_{i,t-1|t-2}\right)\right] \\
&= (1-G)\left(\rho^h z_{it|t-1} - \rho^{h+1} z_{i,t-1|t-2}\right) + G\rho^h\left[\left(z_t - \rho z_{t-1}\right) + \left(\omega_{it} - \rho\omega_{it-1}\right)\right] \\
&= (1-G)\left(z_{i,t+h|t-1} - z_{i,t+h|t-2}\right) + G\rho^h\left[v_t + \omega_{it} - \rho\omega_{it-1}\right].
\end{aligned}$$

Again, averaging across agents, we obtain the analogous relationship for the consensus forecast:

$$\overline{z_{t+h|t}} - \overline{z_{t+h|t-1}} = (1-G)\left(\overline{z_{t+h|t-1}} - \overline{z_{t+h|t-2}}\right) + G\rho^h v_t. \tag{B.2}$$

B.2 Estimation of the model relationships

The equations describing the consensus forecast can be estimated by OLS, because the error terms contain only information that is orthogonal to the explanatory variables. Hence, the OLS β coefficients from (B.1) and (B.2) consistently estimate $(1-G)/G$ and $1-G$, respectively. In contrast, the error terms in the individual regressions are correlated with the right-hand side variables, and thus the OLS estimates are converging to other values. In fact, if the forecasters were forming rational expectations, they would choose a G that ensures that the OLS β is zero in both individual regressions. Since the reduced form regressions reject the null of rational expectations, it is instructive to work out the bias term.

The standard OLS bias in the individual forecast error-forecast revision regression is given by

$$Bias = \frac{Cov\left(z_{i,t+h|t} - z_{i,t+h|t-1}, -\rho^h\omega_{it} + v_{t+h,t}\right)}{Var\left(z_{i,t+h|t} - z_{i,t+h|t-1}\right)}.$$

Using that v_t is i.i.d. over time and ω_{it} is i.i.d. both across time and in the cross-section, we can write

$$\begin{aligned}
Cov\left(z_{i,t+h|t} - z_{i,t+h|t-1}, -\rho^h\omega_{it} + v_{t+h,t}\right) &= Cov\left(\rho^h z_{it|t}, -\rho^h\omega_{it}\right) \\
&= -\rho^{2h} Cov\left(G\left(z_t + \omega_{it}\right) + (1-G)\, z_{it|t-1}, \omega_{it}\right) \\
&= -\rho^{2h} G\sigma_\omega^2.
\end{aligned}$$

For the variance term, we can write

$$\begin{aligned}
Var\left(z_{i,t+h|t} - z_{i,t+h|t-1}\right) &= Var\left(\rho^h z_{it|t} - \rho^h z_{it|t-1}\right) \\
&= \rho^{2h} Var\left[G\left(z_t + \omega_{it} - z_{it|t-1}\right)\right] \\
&= \rho^{2h} G^2 \left[\frac{\sigma_v^2}{1-\rho^2} + \sigma_\omega^2 + \underbrace{Var\left(z_{it|t-1}\right)}_{A} - 2\underbrace{Cov\left(z_t, z_{it|t-1}\right)}_{B}\right],
\end{aligned}$$

where we used that z_t is AR(1). Next, I derive the A and B terms from above:

$$B = Cov\left(\rho z_{t-1}, \rho z_{i,t-1|t-1}\right) = \rho^2 Cov\left(z_t, z_{it|t}\right) = \rho^2 Cov\left[z_t, G\left(z_t+\omega_{it}\right)+\left(1-G\right)z_{it|t-1}\right]$$
$$= \rho^2\left[G\frac{\sigma_v^2}{1-\rho^2}+\left(1-G\right)\underbrace{Cov\left(z_t, z_{it|t-1}\right)}_{B}\right] \implies B = \frac{\rho^2 G\frac{\sigma_v^2}{1-\rho^2}}{1-\left(1-G\right)\rho^2}$$

and similarly

$$A = \rho^2 Var\left(z_{it|t}\right) = \rho^2 Var\left[G\left(z_t+\omega_{it}\right)+\left(1-G\right)z_{it|t-1}\right]$$
$$= \rho^2\left[G^2\left(\frac{\sigma_v^2}{1-\rho^2}+\sigma_\omega^2\right)+\left(1-G\right)^2\underbrace{Var\left(z_{it|t-1}\right)}_{A}+2G\left(1-G\right)\underbrace{Cov\left(z_t, z_{it|t-1}\right)}_{B}\right]$$
$$\implies A = \rho^2\frac{G^2\left(\frac{\sigma_v^2}{1-\rho^2}+\sigma_\omega^2\right)+2G\left(1-G\right)B}{1-\left(1-G\right)^2\rho^2}.$$

Combining our results, we obtain

$$Bias = \frac{-\sigma_\omega^2}{G\left[\frac{\sigma_v^2}{1-\rho^2}+\sigma_\omega^2+A-2B\right]}.$$

It is easy to verify that this expression depends only on the noise-to-signal ratio, and not σ_ω^2 and σ_v^2 separately. Hence, we demonstrated that in the individual regression

$$\widehat{\beta}_{OLS} \xrightarrow{p} \frac{1-G}{G} + Bias\left(G,\rho,\sigma_\omega^2/\sigma_v^2\right) = f(G,\rho,\sigma_\omega^2/\sigma_v^2),$$

as claimed in the text. The persistence parameter ρ can be estimated from the data. Hence, given an estimate of G from (B.1), the individual regression allows us to recover the noise-to-signal ratio that is consistent with the model.

For the alternative individual regression based on current and past forecast revisions, we can write the bias term as follows:

$$\overline{Bias} = \frac{Cov\left(z_{i,t+h|t-1}-z_{i,t+h|t-2}, G\rho^h\left[v_t+\omega_{it}-\rho\omega_{it-1}\right]\right)}{Var\left(z_{i,t+h|t-1}-z_{i,t+h|t-2}\right)} =$$
$$= G\frac{Cov\left(\rho z_{i,t+h-1|t-1}-\rho z_{i,t+h-1|t-2}, -\rho^{h+1}\omega_{it-1}\right)}{Var\left(\rho z_{i,t+h-1|t-1}-\rho z_{i,t+h-1|t-2}\right)}$$
$$= G\frac{Cov\left(z_{i,t-1+h|t-1}-z_{i,t-1+h|t-2}, -\rho^h\omega_{it-1}\right)}{Var\left(z_{i,t-1+h|t-1}-z_{i,t-1+h|t-2}\right)} = G * Bias.$$

Hence, we proved the claim in the text that

$$\widehat{\overline{\beta}}_{OLS} \xrightarrow{p} (1-G) + \overline{Bias}\left(G,\rho,\sigma_\omega^2/\sigma_v^2\right) = G\widehat{\beta}_{OLS}.$$

B.3 Signal noise and forecaster disagreement in the model

In the model, the only source of heterogeneity is agents' private information. Thus, there is a direct relationship between the idiosyncratic noise in private signals and cross-sectional disagreement. The cross-sectional dispersion of h-step ahead forecasts reported at time t can be written as

$$\begin{aligned} Var_i\left(z_{i,t+h|t}\right) &= \rho^{2h} Var_i\left(z_{it|t}\right) = \rho^{2h} Var_i\left[G\left(z_t + \omega_{it}\right) + (1-G)\, z_{it|t-1}\right] \\ &= \rho^{2h}\left[G^2\sigma_\omega^2 + (1-G)^2\, Var_i\left(z_{it|t-1}\right)\right] \\ &= \rho^{2h}\left[G^2\sigma_\omega^2 + (1-G)^2\,\rho^2 Var_i\left(z_{it|t}\right)\right] \\ &\implies Var_i\left(z_{i,t+h|t}\right) = \rho^{2h}\frac{G^2\sigma_\omega^2}{1-(1-G)^2\rho^2}. \end{aligned}$$

Appendix C

Appendix to Chapter 3

C.1 Details of data construction

This appendix describes in detail the data source and the manipulations that were carried out before the statistical analysis.

All data comes from the Bureau van Dijk–Bankscope database. Along with bank characteristics, I retrieved all available data between 1991 and 2010 for the following variables: Total Assets, Total Earning assets, Equity, Equity/Total assets, Net interest margin. I converted all nominal figures into million USD using the exchange rate corresponding to the statement date. In some non-European countries, the annual filing date is not 31 December. I assigned each statement to the year whose end is closer to the filing date. For example, a statement filed in March 2000 would be assigned to 1999, while a statement filed in October 2000 would be assigned to 2000. After these initial steps, I used the following filters:

- I kept observations only between 1993 and 2009 due to low coverage in other years.
- I kept only countries listed in Table C.1.
- I kept only institutions categorized as commercial banks, savings banks or co-operative banks. This means dropping real estate/mortgage banks, medium and long term credit banks, investment banks/securities houses, Islamic banks, non banking credit institutions, specialized governmental credit institutions, bank holdings and holding companies, central banks, and multi-lateral governmental banks.
- I deleted a very small number of nonsense observations, such as equity ratios with negative or above 100% values.

The next and most time consuming task was to eliminate duplicated information resulting from multiple statements. In most instances, Bankscope does not provide a single unique statement per bank over the entire sample period. Major changes in the accounting practices, e.g., the switch to consolidated financial statements, are accompanied by the introduction of a new separate time series for an entity. Therefore, several different financial statements per entity may be available for a given reporting period, e.g., each representing a different basis of consolidation. This requires defining rules for selecting and merging these statements to obtain one unique time series per bank entity.

My primary goal was to use the most aggregated statement available in each year. If both consolidated and unconsolidated statements have been filed in a given year, the consolidated statement is considered as the more aggregate version. If in given year or over the entire sample period only unconsolidated statements are available, I assume that they represent the most senior information available.

In practice, I used Bankscope's consolidation codes to apply the above principle. Bankscope has six codes for consolidation (C2, C1, C* and U2, U1, U*), where C indicates a consolidated and U denotes an unconsolidated statement. The extension "2" indicates that both a consolidated and an unconsolidated statement exist for a bank (codes C2 and U2) at some point of time. Accordingly, the codes C1 and U1 indicate that no companion statement exists. C* and U* indicate that additional statements have been filed. This leads to the following seniority ranking of statements filed (assuming that consolidated statements represent the most senior information available): C2/C1>C*>U1>U*>U2.

Table C.2 contains the number of observations and median values for various variables for each country. Once I cleaned the micro data, I aggregated up the variables to the country level. All country level ratios are weighted averages, so for example the equity-to-assets ratio in a country in a given year is the asset-weighted average of individual bank equity ratios in that year.

C.2 Comparative statics in the model

In the text, I gave an intuitive argument why lower charter values can lead to higher debt ratios, and illustrated with numerical examples that my simple model is able to deliver this prediction. Here, I put more structure on the problem and derive sufficient conditions under which the right comparative statics obtain.[1]

Let R be a continuous random variable with cumulative distribution function F. As a realistic assumption, I impose $R > 0$ for all realizations; i.e. the bank cannot have negative *gross* return on its investment. The return can get arbitrarily small, but even if the whole loan portfolio becomes non-performing, its value cannot become negative. Formally, $F(0) = 0$. I also assume that $F'(0) = 0$. This means, loosely speaking, that returns very close to zero are also very unlikely.

Claim 6 *With the above assumptions, the bank never chooses full equity financing.*

Proof. The first order condition of the problem in equation (3.2) in the text is

$$-F'(d)[\gamma d + x] + [1 - F(d)]\gamma = 0. \tag{C.1}$$

[1]These conditions are by no means necessary, as my numerical simulations (not reported here) revealed.

With our assumptions, it is easy to check that at $d = 0$ the left-hand side is equal to $\gamma > 0$, so the bank has an incentive to increase the debt ratio above zero. ■

To ensure that $d < d_{max}$, the left-hand side of equation (C.1) must be negative at $d = d_{max}$. Intuitively, we need that the threat of default be strong enough. If the distribution of R is such that there is no downside risk at all, then the bank will increase d to its maximum level, because it can collect the tax refund without jeopardizing losing the charter value. Since the general conditions are not particularly enlightening, I just assume that F is not tilted towards high values too much, so the problem admits an interior optimum.

The next lemma derives a sufficient condition that ensures that my main comparative statics result hold.

Claim 7 *If $F''(d) \geq 0$ for all $0 < d < d_{max}$, then reducing the charter value increases the optimal debt ratio.*

Proof. Denote the optimal debt ratio by d^*. Applying the implicit function theorem to equation (C.1), we obtain

$$\frac{\partial d^*}{\partial x} = -\frac{\overbrace{F'(d^*)}^{+}}{\underbrace{2\gamma F'(d^*)}_{+} + \underbrace{[\gamma d^* + x]}_{+} F''(d^*)}$$

The result follows from the sign of the derivative. ■

How can one interpret the $F''(d) \geq 0$ condition? The easiest way to visualize the intuition is in the case of unimodal distributions. The condition says that for $0 < d < d_{max}$ we are still to the left of the mode, so the pdf is still increasing. In my model the risk free interest rate is 0, so a very realistic assumption is that $E[R] > 1$, i.e. the risky investment carries a risk premium. If the return distribution is not too skewed, then the mode is close the expected value, so for any $d < 1$ we have $F''(d) \geq 0$.[2]

[2]Notice, however, that the condition is not necessary. The other positive terms can easily dominate F'' even to the right of the mode.

Table C.1: List of sample countries

Description	Country	EU/EEA	Euro
Treatment group EU members who adopted the euro in 1999 and Greece	Austria	1995	1999
	Belgium	1957	1999
	Finland	1995	1999
	France	1957	1999
	Germany	1957	1999
	Ireland	1973	1999
	Italy	1957	1999
	Luxembourg	1957	1999
	Netherlands	1957	1999
	Portugal	1986	1999
	Spain	1986	1999
	Greece	1981	2001
Control I (Old EU) Old EU/EEA members who did not adopt the euro	United Kingdom	1973	X
	Sweden	1995	X
	Norway	1995	X
Control II (Non EU) Non-EU developed countries	Australia	X	X
	Canada	X	X
	Japan	X	X
	Korea	X	X
	New Zealand	X	X
	Switzerland	X	X
	United States	X	X
Control III (New EU) New EU members	Cyprus	2004	2008
	Czech Republic	2004	X
	Estonia	2004	2011
	Hungary	2004	X
	Latvia	2004	X
	Lithuania	2004	X
	Malta	2004	2008
	Poland	2004	X
	Slovakia	2004	2009
	Slovenia	2004	2007

Note: The table shows the classification of countries into treatment and control groups. The third column provides the year of entry into the European Union/European Economic Area. The last column is the year of introducing the euro in that country. An “X” indicates a country with its own currency.

Table C.2: Summary statistics (medians)

Country	Obs. (Bank-year)	Total assets (million USD)	Earning assets (million USD)	Equity to total assets (%)	Net interest margin (%)
Austria	1254	328.89	306.45	6.2	2.8
Belgium	553	833.44	803.92	6.3	1.9
Finland	63	12339.28	11143.09	5.5	1.9
France	2700	1896.01	1787.32	6.7	2.4
Germany	14074	420.12	399.19	4.9	2.9
Greece	144	1883.13	1670.20	8.9	3.0
Ireland	246	3092.60	3011.54	6.0	1.0
Italy	5013	206.04	191.29	11.6	3.7
Luxembourg	861	1115.51	1066.80	3.9	0.8
Netherlands	348	1978.67	1847.59	6.8	1.7
Portugal	227	2890.58	2603.98	6.5	2.6
Spain	1204	1803.17	1675.27	7.6	3.3
United Kingdom	1192	1066.85	995.30	8.6	2.0
Sweden	356	145.34	137.15	12.3	3.8
Norway	357	857.38	819.61	8.1	3.0
Australia	227	3098.52	3025.64	6.3	2.2
Canada	399	836.71	757.57	7.1	2.1
Japan	4642	1563.09	1488.47	4.9	2.1
Korea	183	13166.80	10959.04	4.2	2.3
New Zealand	67	10965.32	10064.90	4.9	2.6
Switzerland	2088	179.63	170.36	7.3	1.7
United States	47856	97.87	89.80	9.5	4.1
Cyprus	123	321.32	305.43	6.7	2.9
Czech Republic	194	783.05	710.11	6.8	2.4
Estonia	50	203.32	159.20	9.9	5.6
Hungary	193	581.08	504.02	9.3	4.3
Latvia	170	103.12	85.69	10.3	4.5
Lithuania	75	138.59	112.00	12.6	5.1
Malta	60	413.29	400.65	7.0	2.3
Poland	333	463.51	432.37	10.7	5.3
Slovakia	135	566.50	523.10	8.5	3.3
Slovenia	168	294.76	274.69	11.4	3.8
All	85555	192.58	178.60	8.5	3.6

Note: The table shows median values for all bank-year observations in each country from 1993-2009.