INFORMATION FRICTIONS IN MACRO-FINANCE

LUCA GEMMI

A dissertation

submitted to the Faculty of

the department of Economics

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

March, 2022

INFORMATION FRICTIONS IN MACRO-FINANCE

LUCA GEMMI

Advisors: Associate Professor Ryan Chahrour, Assistant Professor Jaromir Nosal, and Assistant Professor Rosen Valchev

Abstract

I study how economic conditions and strategic incentives affect belief formation of rational agents with a limited information processing capacity. I study the impact of cognitive and information frictions on individual risk taking, investment and portfolio choice, and their implications on aggregate macroeconomic fluctuations.

In my first chapter *"Rational Overoptimism and Moral hazard in Credit Booms"* I develop a framework in which over optimism in credit booms originates from rational decisions of managers. Because of moral hazard, managers pay too little attention to the aggregate conditions that generate risk, leading them to over borrow and over invest during booms. Periods of low risk premia predict higher default rates, higher probability of crises and systematic negative banks excess returns, in line with existing evidence. I document a negative relation between the convexity of CEO's compensation and their information on a larger sample of firms, which is consistent with my theory. My model implies that compensation regulation can play an important role in macro prudential policy.

In my second chapter *"Biased Surveys"* Rosen Valchev and I improve on the standard tests for the FIRE hypothesis by allowing for both public and private information, and find new interesting results. First, we propose a new empirical strategy that can accommodate this richer information structure, and find that the true degree of information rigidity is about a third higher than previously estimated. Second, we find that individual forecasts over-react to private information but under-

react to public information. We show that this is consistent with a theory of strategic diversification incentives in forecast reporting, where forecasters are rational but report a biased measure of their true expectations. This has two effects. First, it generates what looks like behavioral "over-reaction" in expectations, and second biases the information rigidity estimate further downward. Overall, our results caution against the use of survey of forecasts as a direct measure of expectations, and suggest that the true underlying beliefs are rational, but suffer from a much larger degree of imperfect information than previously thought. This has particularly profound implications for monetary policy, where inflation expectations play a key role.

I explore further how economic incentives shape beliefs in my third chapter *"International Trade and Portfolio Diversification"*. I show that information choice can explain the puzzling positive relation between bilateral investment and trade across countries. I present a model of endogenous information with both investment in assets and income from trade. While standard model of risk-hedging would require agents to invest in non-trading countries to diversify income risk, I show that limited information capacity and preferences for early resolution of uncertainty reverse this result. The intuition is that investors collect more information on trading partners to reduce income uncertainty, and therefore perceive their equity as less risky. I find that allowing for information choice reduces the role of risk hedging on portfolio decisions. I test my model's implied relation between trade and attention in the data and find robust empirical support.

Contents

List of Tables

List of Figures

Chapter 1

Rational Overoptimism and Moral Hazard in Credit Booms

1.1 Introduction

Recent empirical works has revived the longstanding hypothesis that boom-and-bust credit cycles are driven by overoptimistic beliefs (Minsky, 1977; Kindleberger, 1978). In particular, empirical evidence documents that high credit growth and low risk premia significantly predict subsequent financial crises (Schularick and Taylor, 2012; Jordà et al., 2013; Krishnamurthy and Muir, 2017; Greenwood et al., 2020). Two additional facts point towards overoptimistic beliefs as an explanation for this evidence. First, credit booms also predict low and even negative excess returns on bank stocks (Baron and Xiong, 2017). Second, forecasts are systematically too optimistic when credit spreads are low (Bordalo et al., 2018b; Gulen et al., 2019). Behavioral models of extrapolative beliefs have been particularly successful in explaining such systematic bias in belief formation (Maxted, 2019; Bordalo et al., 2021; Krishnamurthy and Li, 2021).

While existing theories of overoptimism preserve full information and depart from rational expectations, I provide evidence on the importance of information frictions during booms. I compare actual real GDP growth with survey forecasts during booms and NBER recessions, and show that panelists consistently underestimate real output in booms and overestimate it in recessions. I document a similar pattern of belief underestimation in housing starts growth during the housing bubble that preceded the financial crisis of 2008-2009. This evidence on systematic belief under-reaction is consistent with imperfect information about aggregate quantities and in line with a recent literature on information dispersion (Coibion and Gorodnichenko, 2015; Coibion et al., 2018; Gemmi and Valchev, 2021).

First, I develop a theory of credit booms where overoptimism originates from rational inattention to aggregate risk factors. In my model credit, an aggregate productivity shock leads to an increase in borrowing and production of firms who face the same downward sloping demand for their combined output. Because higher aggregate production implies lower selling price, inattention to competitors' investment decisions cause firms to form overoptimistic expectations about their own revenues. Inattentive firms over borrow and over invest, causing an excess supply in the good market which further amplifies the decline in price. As firms' revenues are lower than expected, their default risk increases. My model implies that even fully rational agents can be systematically overoptimistic in credit booms and overpessimistic in busts. Moreover, because inattentive banks underestimate borrower's probability of default, they misprice risk and register negative excess returns after credit booms, consistently with the existing evidence.

Second, I show that inattention to risk factors can be ascribed to moral hazard incentives in information choice. Because managers with convex compensation structures are less exposed to company's losses, they have a lower marginal benefit of information, resulting in lower attention to aggregate conditions. Uninformed managers underestimate the increase in competition and decline in revenues after booms and are overoptimistic about their company's revenues. As a result, moral hazard incentives don't just lead to excessive risk taking given beliefs, but also inattention to risk and overoptimistic beliefs in boom periods. This result helps connect the two narratives of excess risk taking before the financial crisis of 2008-2009: the initial criticisms toward managers' moral hazard incentives (e.g. Blinder 2009) and the following behavioral overoptimism view (e.g., Gennaioli and Shleifer 2018). I show that overoptimism is in fact a consequence of moral hazard incentives.

Finally, I provide empirical evidence on the relation between manager's compensation and information choice on a large sample of US firms. I look at the relation between firm's CEO compensation and its earning guidance and document that higher compensation asymmetry, measured as share of stock options for a given stock of shares, is positively correlated with inattention, measured as squared forecast errors on future profits. The evidence documents a negative relation between moral hazard incentive in information choice, consistently with my model.

Because beliefs are rational, my model implies that policy makers can reduce overoptimism in credit booms by regulating manager's incentives to collect information. Informed managers reduce borrowing and investment in credit booms, mitigating economic fluctuations. However information provision through public announcement or direct communication would still be costly for managers to process. Instead, solving the moral hazard by regulating managers' compensation would not only solve their excess risk taking in investment, but also encourage them to pay attention to aggregate risk factors.

Model I embed compensation incentives and information choice in a macroeconomic model with endogenous default. The model features a continuum of bank-firm pairs, which I refer as islands. Firms demand loans from banks in order to finance investment, while banks get funding at a constant risk free rate on international markets. Firms and banks are run by managers with a convex compensation scheme.

I introduce two important elements to an otherwise standard setting. First, strategic substitutability between islands. I assume each firm produces intermediate goods, which are acquired by a unique aggregate final good producer with downward sloping demand. Aggregate credit booms lead to an increase in aggregate supply of intermediate goods, which lowers the individual firm's selling price and therefore its revenues. Second, I introduce incomplete information. Following the Lucas island framework, I assume agents are not able to freely observe aggregate prices and quantities.[1] However, I allow bank and firm managers on each island to pay an information cost to observe aggregate shocks and therefore investment decisions of competitors.[2]

Firm's productivity depends on local and aggregate shocks but, because of the competition in the intermediate goods market, firms benefit more from local than aggregate shocks. Local shocks improve firm's fundamentals and reduce its default probability, resulting in higher equilibrium debt and lower spreads. On the other hand, aggregate shocks also increase production of competitors and therefore lower firm's expected revenue and increase its default probability relative to a local shock with the same magnitude. While the first effect is standard in the literature that abstracts from competition between islands, the second effect is novel and implies a strategic interaction between islands.

First, I show that the full information model is not able to qualitatively match the existing evidence on risk premia in a credit boom. If managers observe aggregate shocks, the economy is always safer in credit booms, which implies lower risk premia. Even if the negative price externality has a dampening effect on the credit boom, the model is qualitative similar to a standard model without this additional channel (Strebulaev and Whited, 2011). Because the economy is safer after a boom,

[1] This assumption is consistent with the decentralized nature of bank credit market.

[2] I follow the rational inattention literature (Sims, 2003, 2006) in interpreting the information cost as a cognitive cost agents pay in order to processing information which could be freely accessible.

the model does not match the existing evidence.

I show that the model with dispersed information is instead able to match the existing evidence on credit cycles. If managers do not observe aggregate shocks, they incorrectly attribute the boom primarily to a local shock and underestimate the increase in production of competitors. As a result, they over-borrow and over-invest, further overheating the economy. Even if perceived risk and risk premia decline, default rate increases. The model is consistent with the existing empirical evidence. First, credit growth predicts higher average probability of default (Krishnamurthy and Muir, 2017). Second, low risk premia also predict higher average probability of default (Krishnamurthy and Muir, 2017). Third, bank's excess return during the boom and bust is negative on average (Baron and Xiong, 2017).

Next, I endogenize information and show that moral hazard incentives discourage information acquisition. Because managers with convex compensation structures are less exposed to company's losses, their marginal benefit of information is lower and they decide to collect less information. As a result, they will be inattentive to the endogenous increase in risk during credit booms. Importantly, the excess risk taking in booms depends on managers' inattention to risk and not simply on higher risk taking in investment choice. In order to isolate the information channel of moral hazard, I shut down information choice and allow managers to observe aggregates. I show that standard compensation risk taking incentives alone without information choice are not able to qualitatively match the data.

Finally, I embed the model in a infinite-period framework to study its implication for credit cycles and relate it to the existing evidence. I show that the model with a realistic calibration is able to reproduce two important sets of moments in the data. First, my model matches the systematic decrease in spreads and increase in credit growth before financial crises. Second, it reproduces the predictive power of decline in spreads and increase in credit in forecasting financial crises.

Empirics I find empirical evidence on the model's implied positive relation between CEO's compensation convexity and squared forecast errors. I measure CEOs' beliefs with firm's forecasts on future earnings per share from the *IBES Guidance* database.[3] I measure CEOs' compensation convexity as options stock holding controlling for equity shares holding and additional CEO and firm controls. I find that higher compensation convexity is associated with larger manager's squared forecast errors, in line with the model's implication.

Contribution to the literature This paper contributes to several strands of the literature. First, the growing body of research about credit cycles. In addition to the already mentioned empirical work (Schularick and Taylor, 2012; Jordà et al., 2013; Krishnamurthy and Muir, 2017; Baron and Xiong, 2017; López-Salido et al., 2017; Mian et al., 2017; Greenwood et al., 2020), this paper relates to the theoretical research on financial crises, which can be divided in two categories. The first emphasizes the role of behavioral bias in belief formation and credit market sentiments (Bordalo et al., 2018b; Greenwood et al., 2019; Maxted, 2019; Farhi and Werning, 2020). The most related is Bordalo et al. (2021), which embeds extrapolative expectations in a firm dynamic model with lending and default. In their model, beliefs overreact to good news, leading to overoptimism in credit booms. In my model overoptimism originates instead from underreaction to bad news. As a result, forecast errors exhibits predictability even in a fully rational setting.

A second line of research emphasizes the role of financial frictions in intermediation as sources of fragility (Gertler and Kiyotaki, 2010; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2019; Jeanne and Korinek, 2019; Bianchi and Mendoza, 2020). This class of models use full information and strategic complementarity in leverage choices to rationalize the overaccumulation of debt during booms, as individuals do not internalize the exter-

[3] The underlying assumption is that the earning projection released by the firm, even if not personally computed by the CEO, has been approved by him (Otto, 2014).

nality effects of their decision on the whole economy. Differently from them, in my model financial fragility originates from strategic substitutability and incomplete information. If agents knew about the increase in aggregate risk, they would reduce leverage and therefore reduce risk. In fact, in my model financial fragility increases because managers do not pay attention to it. As a result, while Fisherian models exhibit cooperation among investors who ride the bubble as long as other ride it, my model exhibits competition between investors, as they want to exit the bubble before it burst.

My paper also relates to the literature on strategic games with incomplete information (Woodford, 2001; Coibion and Gorodnichenko, 2012; Maćkowiak and Wiederholt, 2015). While dispersed information and strategic substitutability lead to amplification of partial equilibrium effects as in Angeletos and Lian (2017), I study its implication for pricing of risk in credit booms. Similarly to Kohlhas and Walther (2020), agents here pay asymmetric attention to local and aggregate quantities, which leads to "extrapolative beliefs" even in a rational setting. Differently from them, the determinant of the attention allocation is not the difference in shock volatility, but moral hazard incentives.

Finally, this paper contributes to the literature on compensation incentives. In addition to the large body of research on CEO compensation (see Edmans et al. 2017 for a review), I mostly relate to the works studying the impact of compensation on information. Mackowiak and Wiederholt (2012) show that limited liability reduces optimal information choice, while Lindbeck and Weibull (2017) study optimal contracts between principal and manager in rational inattention setting. Differently form them, this paper abstracts from optimal contracts, but contributes by documenting the impact of compensation on information in the data and studying its implication on credit cycles. My empirical results are complementary to Cole et al. (2014), which provides experimental evidence on the impact of compensation on loan officers' screening effort.

1.2 Motivational Evidence on Beliefs in Booms

While existing theories of overoptimism preserve full information and depart from rational expectation, in this section I provide evidence which points towards the importance of information frictions in business cycles. In particular, I document that aggregate beliefs under-react to changes in macroeconomic quantities in booms and busts, consistent with models of dispersed information.[4]

First, I look at business cycle frequency fluctuations of forecast errors on real GDP growth by comparing the average errors in booms and recessions. Forecast errors are defined as $fe_t = x_t - f_t(x_t)$, where x_t is the average annualized growth of real GDP in the current and the next three quarters, and $f_t(x_t)$ the average (consensus) forecast in quarter t about annualized growth of real GDP at the same horizon. Forecast data are from the Survey of Professional Forecasters, and a positive forecast errors imply that the consensus forecast underestimate the actual GDP growth. Figure 1.1 shows that forecasters underestimate real output during booms and overestimate them during NBER recessions. This evidence suggests that at the aggregate level expectations display underreaction, and not overreaction, to changes in macroeconomic quantities.

In addition to the business cycle frequency, I provide evidence for belief underreaction in the most recent credit boom-and-bust episode. Financial crises are less frequent than business cycle recession, and given the limited time span of expectations data the only meaningful credit boom-and-bust I can consider is the recent financial crisis of 2007-2008. Figure 1.2 plots annualized growth forecasts and realizations of housing starts, averaged across the current and the next three quar-

[4] A leading behavioral theory of overoptimism is belief extrapolation, and in particular diagnostic expectations, which causes agents to over-react to recent news (Gennaioli and Shleifer, 2010; Bordalo et al., 2018b, 2021).

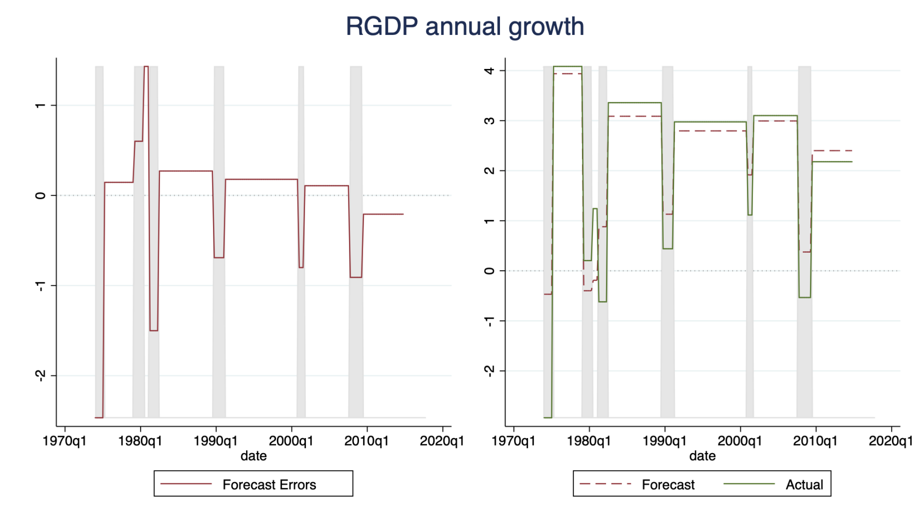


Figure 1.1: Forecast errors on Real GDP growth

Notes: Left panel: the red line plots the forecast errors on annualized real GDP growth averaged between shaded area. Forecast errors are defined as $fe_t = x_t - f_t(x_t)$, where x_t is the average annualized growth of real GDP in the current and the next three quarters, and $f_t(x_t)$ the average (consensus) forecast in quarter t about annualized growth of real GDP in the current and the next three quarters. The shaded area indicates the NBER recession dates. Right panel: the dashed red line plots the average forecast on annualized real GDP growth $f_t(x_t)$, while the solid green line the actual real GDP growth x_t. All expectation data are from the Survey of Professional Forecasters, collected by the Federal reserve's Bank of Philadelphia

ters. The pattern is similar to the previous figure and it suggests that forecasters underestimated housing starts growth during the boom. In the next section I show how underestimation of an increase in supply leads to overestimation of the equilibrium market price, which might shed some light on the apparent overoptimism that boosted the housing bubble in the years preceding the crisis.

In addition to the evidence reported here, a growing literature employs surveys of professional forecasters to document the importance of information frictions against the full information hypothesis (Coibion and Gorodnichenko, 2012, 2015; Gemmi and Valchev, 2021).[5] The evidence of aggregate stickiness in belief updating supports model of dispersed information, where agents have access to different in-

[5] Bordalo et al. (2018a) provides evidence supporting behavioral overreaction in survey individual-level forecasts on financial and macroeconomic variables. However, they still find dispersed information and belief stickiness at the consensus level. Moreover, Gemmi and Valchev (2021) provide further evidence on survey individual forecast which are inconsistent with the diagnostic expectation framework.

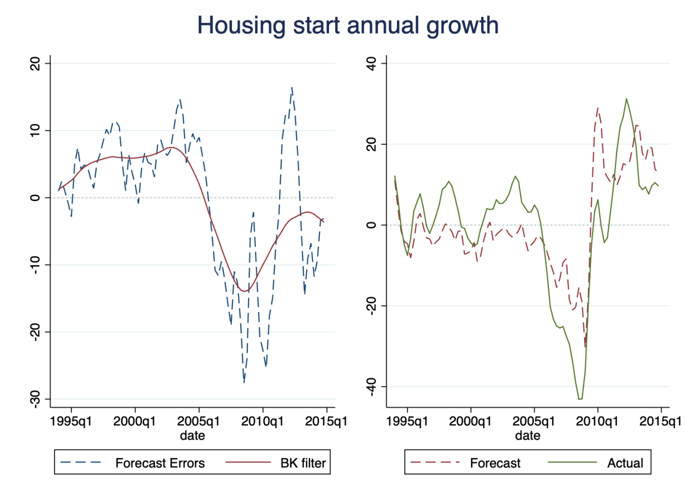


Figure 1.2: Forecast errors on Housing Start

Notes: The blue line plots the forecast errors on annualized housing start growth from the Survey of Professional Forecasters, collected by the Federal reserve's Bank of Philadelphia. Forecast errors are defined as $fe_t = x_t - f_t(x_t)$, where x_t is the average annualized growth of housing starts in the current and the next three quarters, and $f_t(x_t)$ the average (consensus) forecast in quarter t about annualized growth of housing starts in the current and the next three quarters. The red line plot the Baxter-King filtered trend, where I filtered out periods lower than 32.

formation and are always in disagreement about the fundamentals. Moreover, the professional forecaster's expectations data I use here are likely to underestimate the amount of information friction of firms. In line with this, Coibion et al. (2018) study firm's level expectation and find stronger results. Managers' expectations display much more disagreement than professional forecasters, and this disagreement applies to both future and current economic condition. Moreover, they find that their belief updating is consistent with the Bayesian framework and their attention allocation to aggregates depends on incentives.

In summary, the evidence on aggregate expectations are consistent with information frictions that hinder the diffusion of information or the incorporation of new information in agent's beliefs (Sims, 2003; Woodford, 2001). In the following section I present a model consistent with the data, where overoptimism originates from incomplete information about aggregate quantities.

1.3 Model of inattentive credit booms

The economy is populated by a continuum of islands $j \in [0,1]$ and each island is populated by a firm-bank pair.[6] Banks in each island collect funds at the risk free rate in international markets and lend to the firm at a premium above the funding rate to cover for repayment risk. Firms borrow from banks in order to finance investment and production of intermediate goods, which they sell to a unique aggregate final good producer. If revenues are higher than outstanding debt, the firm repays the bank and keep the net profit, and otherwise it defaults.

The model is divided in three stages. First, before receiving any information each bank-firm pair decides whether they want to observe aggregate shocks in the next stage. Second, they observe information and bargain on loans and loan rates. Finally, shocks realize and firms repay or default. Rather than a description of business cycles, the model is intended to describe the phases of a financial bubble, with the second stage representing the building up of the bubble and the third stage its burst.

Final good producer The economy features a representative final good producer, acquiring a bundle of intermediate goods $M = \left[\int^{j} M_j^{\xi} dj\right]^{\frac{1}{\xi}}$ with elasticity of substitution $\frac{1}{1-\xi}$, in order to produce final good $Y = M^{\nu}$. Therefore, the demand function for intermediate goods M_j in stage 3 equals:

$$p_j = \nu M^{\nu-\xi} M_j^{\xi-1} \tag{1.1}$$

The demand for intermediate good M_j could increase or decrease in aggregate production M depending on the degree of decreasing return to scale in final good

[6] The island assumption reflects the importance of banking relationship and the cost faced by borrowers in switching lender (Chodorow-Reich, 2014). I assume that the sorting of lenders and borrowers across island happens before markets open and information is observed, when there is no heterogeneity in firms and banks characteristics.

production and the elasticity of substitution between goods. If $\nu < \xi$, higher aggregate supply of intermediates M lead to lower price p_j and therefore lower revenues for intermediate producer j. Conversely, if $\nu < \xi$, higher aggregate supply of intermediates M lead to higher price p_j and therefore higher revenues for intermediate producer j. opposite. I assume $\nu < \xi$ and in section 1.4.2 I show that this condition holds under fairly mild assumptions, such as an equal markup in intermediate and final good sectors.

Firms In the second stage, firms in island j borrows b_j from the bank in order to purchase capital inputs and cover the capital adjustment cost. For simplicity, I assume firms start with zero net worth and therefore borrowing equal $b_j = k_j + \phi\frac{k_j^2}{2}$. In the third stage, firms combine labor l_j, pre-installed capital k_j and productivity A_j with production function

$$M_j = A_j^{\zeta} k_j{}^{\tilde{\alpha}} l_j^{1-\tilde{\alpha}} \tag{1.2}$$

The parameter $\tilde{\alpha} \in (0,1)$ represents the capital share. Firms hire labor in the third stage after observing the shocks realization and pay workers before repaying their debt to the bank. Define the operating profits of the firm as $\pi_j = p_j M_j - w l_j$. We can maximize labor out of the problem and substitute for the demand function (1.1) to obtain net operating profit as function of only capital, technology and aggregate supply of intermediates

$$\pi(A_j, k_j, M) = \Lambda(M) A_j k_j^{\alpha} \tag{1.3}$$

where $\alpha = \frac{\tilde{\alpha}\xi}{1-(1-\tilde{\alpha})\xi}$, $\Lambda(M) = \nu^{\frac{1}{1-(1-\alpha)\xi}} M^{\frac{\nu-\xi}{1-(1-\alpha)\xi}}$

$$M = \left\{ \left[\frac{w}{(1-\alpha)\xi\nu} \right]^{\frac{(1-\alpha)}{(1-\alpha)\xi-1}} \left[\int^{N} A_j k_j^{\alpha} dj \right]^{\frac{1}{\xi}} \right\}^{\frac{1-(1-\alpha)\xi}{1-(1-\alpha)\nu}} \tag{1.4}$$

Here I have normalized the parameter ζ so that the profit function is linear in technology and the real wage w so that the constant multiplying $\Lambda(M)$ in the profit function equals 1.

Firms payoff in stage 3 are as follows:

$$d_{firm,j} = \begin{cases} (1-\tau)[\pi(A_j, k_j, M) - (1+r_j)b_j] & \text{if } \pi(A_j, k_j, M) \geq (1+r_j)b_j \\ -c_d k_j, & \text{if } \pi(A_j, k_j, M) < (1+r_j)b_j \end{cases} \tag{1.5}$$

If profits are larger than the outstanding debt $(1+r_j)b_j$, the firm repays the bank and keep the difference as dividends, minus a tax rate τ. If the profits are not enough to repay the outstanding debt, the firm pays a default cost c_d proportional to installed capital, which can be thought as a liquidation or reorganization cost following the bankruptcy procedure.[7]

Banks Banks in each island j are deep-pocketed and risk-neutral. In the second stage they borrow at risk free rate r^f in the international market to finance the risky loan to firms b_j at loan rate r_j. They maximize their expected profits in the third stage, which equal

$$d_{bank,j} = \begin{cases} [(1+r_j) - (1+r^f)]b_j & \text{if } \pi(A_j, k_j, M) \geq (1+r_j)b_j \\ -(1+r_j)b_j & \text{if } \pi(A_j, k_j, M) < (1+r_j)b_j \end{cases} \tag{1.6}$$

where risk free rate r^f is exogenous and equilibrium loan rate r_j is determined in stage 2. Because firm's revenues are lost when the firm defaults, default is inefficient in this economy.

Exogenous shocks The logarithm of local technology A_j in each island j is the sum of two independent components: an i.i.d. local island component ϵ_j and an

[7] I consider here a form of "reorganization" bankruptcy, as in Chapter 11 of US bankruptcy code, under which the firm is allow to keep operating after a period of reorganization. This procedure implies some cost such as reputation costs, asset fire sales, loss of customer or supplier relationships, legal and accounting fees, and costs of changing management, which I assume depend on the size of the firm (Branch, 2002; Bris et al., 2006).

aggregate component θ:

$$ln(A_j) = \epsilon_j + \theta \tag{1.7}$$

Agents in each island have common prior $\epsilon_j \sim N(0, \sigma_\epsilon^2)$ and $\theta \sim N(0, \sigma_\theta^2)$. Both shocks realize in stage 3 and determine aggregate and local production.

1.3.1 Firm Manager's compensation

I assume that firms and banks are not run by the shareholders but by risk-neutral managers, who receives a compensation in shares and stock options. In particular, the manager gets $1 - \psi$ shares of company's equity and ψ stock options. I consider manager's compensation convexity as a source of moral hazard incentives in the model for three reasons. First, it is one of the most studied source of moral hazard incentives (Edmans et al., 2017).[8] Second, in the aftermath of the financial crisis of 2008-2009 compensation policies have been suggested as likely culprits for the excessive risk-taking that led to the crisis (e.g. Bebchuk et al. (2010)). Third, I am going to test the model's implications in the data using option holdings in section 2.2.

The manager's compensation structure is as follows:

$$w = \begin{cases} (1-\psi)d_j + \psi(d_j - \tilde{P}) & \text{if } d_j \geq \tilde{P} \\ (1-\psi)d_j & \text{if } d_j < \tilde{P} \end{cases} \tag{1.8}$$

where d_j is the company's payoff, bank or firm, and $\tilde{P}$ is the profit level corresponding to the exercise price of manager's options. Figure 1.3 illustrates the relation between company and manager's payoffs. The larger the amount of options in manager's compensation scheme ψ, the lower is his exposure to company's losses

[8] Stock option compensation in US companies has increased considerably during the 1980s, and especially in the 1990s, becoming the largest component of executive pay. Options increased from only 19% of manager's pay in 1992 to 49% by 2000, and start declining from mid-2000 and in 2014 they represent 16% of the pay (Edmans et al., 2017).

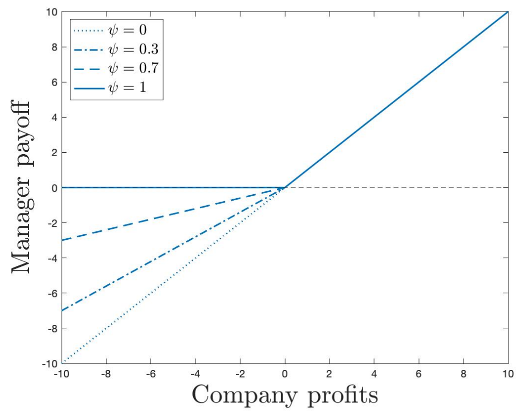


Figure 1.3: Manager's compensation

and therefore higher his insurance against company's losses.[9]

I assume for simplicity $\tilde{P} = 0$, meaning that manager's options are in the money when the profits of the firm are positive, i.e. in the non-default state. Therefore firm manager's final payoff is given by

$$w_{firm,j} = \begin{cases} (1-\tau)[\pi(A_j, k_j, M) - (1+r_j)b_j] & \text{if } \pi(A_j, k_j, M) \geq (1+r_j)b_j \\ -(1-\psi)c_d k_j, & \text{if } \pi(A_j, k_j, M) < (1+r_j)b_j \end{cases} \tag{1.10}$$

Even if here I follow the interpretation of ψ as option share of firm manager, one can also simply interpret it as decrease in firm's default cost c_d. In other words, while I focus on the moral hazard problem between shareholder and manager, one can similarly think about the moral hazard problem between lender and borrower.

[9] A more general compensation structure would consist of β_m shares of company's equity, of which ψ are options.

$$w_j = \begin{cases} \beta_m(1-\psi)d_j + \beta_m\psi(d_j - \tilde{P}) & \text{if } d_j \geq \tilde{P} \\ \beta_m(1-\psi)d_j & \text{if } d_j < \tilde{P} \end{cases} \tag{1.9}$$

The net profits for the shareholder are $(1-\beta_m)d_j$ if profit are positive and $\beta_m\psi d_j$ otherwise. In particular, $\beta_m < 1$ in order to ensure a positive expected leftover profit for the shareholders. However, setting $\beta_m = 1$ does not affect qualitatively the results. Moreover, an additional fixed compensation $\bar{w}$ would not affect the manager's incentives and therefore his decisions.

Bank manager compensation is given by

$$w_{bank,j} = \begin{cases} [(1+r_j) - (1+r^f)]b_j & \text{if } \pi(A_j, k_j, M) \geq (1+r_j)b_j \\ -(1-\psi)(1+r_j)b_j & \text{if } \pi(A_j, k_j, M) < (1+r_j)b_j \end{cases} \tag{1.11}$$

One can also interpret ψ more generally as the degree of limited liability of the bank, which could be due to the share of funds coming from insured deposit versus equity (Dell'Ariccia et al., 2014) or as the share of the loan value covered by government guarantees, a major part of the COVID-19 support packages offered by European governments to companies (OECD, 2020).

1.3.2 Stage 2: lending problem

Before shocks realize and production takes place, bank and firm managers in each island decide loan quantity b_j and rate r_j based on their expectation about profits in stage 3. They share the island's surplus by Nash bargaining with the firm holding all the bargaining power, which implies a zero expected profit condition on the lender in line with the literature (e.g. Strebulaev and Whited (2011)).

Information structure The bank and firm on island j share the same information, as any private information would be perfectly revealed by local prices. Before making their borrowing and lending decision, they receive up to two signals. First, they observe a free noisy signal about local productivity:

$$z_j = ln(A_j) + \eta_j \tag{1.12}$$

with $\eta_j \sim N(0, \sigma_\eta^2)$ and $ln(A_j) = \epsilon_j + \theta$.

Second, they may perfectly observe aggregate productivity. Following the Lucas island setting, I assume managers in each islands do not observe aggregate quantities and prices for free. However, in stage 1 bank and firm managers in island j can decide whether to pay a cost to perfectly observe the aggregate shock θ. Let Ω_j be the (common) stage-2 information set of managers in island j: if they pay the cost in stage 1, $\Omega_j = \{z_j, \theta\}$, otherwise $\Omega_j = \{z_j\}$.

Notice that even in absence of direct observation on aggregate technology, the signal is still informative about θ, since as local technology is the sum of local and aggregate components. The local component ϵ_j acts as a nested noise weakening the inference.

Lending decision The bank's expected excess return is

$$\begin{aligned} E[w_{bank,j}|\Omega_j] =& b_j[1 - p(default_j|\Omega_j)](1 + r_j)b_j \\ & - [(1 - \psi) + \psi[1 - p(default_j|\Omega_j)]](1 + r^f)b_j \end{aligned} \tag{1.13}$$

where the posterior default risk equals

$$p(default_j|\Omega_j) = \int_0^{\infty} \int_{-\infty}^{ln\left(\frac{b_j}{\Lambda(M)k(b_j)^{\alpha}}\right)} f(ln(A_j), M|\Omega_j) dA_j dM \tag{1.14}$$

where Ω_j is the information set of island j, $f(ln(A_j), M|\Omega_j)$ the joint posterior density function of $ln(A_j)$ and M_j and $k(b_j) = \phi^{-1}(\sqrt{1 + 2b_j\phi} - 1)$. The loan rate is implicitly determined by the zero expected profit condition on the bank

$$\frac{1 + r_j}{1 + r^f} = \frac{(1 - \psi) + \psi[1 - p(default_j|\Omega_j)]}{[1 - p(default_j|\Omega_j)]} \tag{1.15}$$

The loan rate is proportional to the perceived probability of default, implying that the risk premium on the loan is only proportional to the perceived risk with no time-varying price of risk.[10] A higher compensation convexity ψ lowers the elasticity of

[10] This result derives from the assumption that the firm retains all the bargaining power, which implies a no expected profits condition for the bank. A non-zero bargaining power on the bank will not change the mechanism of the model, but it will change the determination of the risk premium, which could decline for higher quantity of risk if the price of risk decline as well. See appendix for an alternative calibration of the model where the bank has a non-zero bargaining power.

the spread with respect to perceived risk.

The expected payoff of the firm's manager conditioning on second stage information set is

$$E[w_{firm,j}|\Omega_j] = (1-\tau)\int_0^\infty \int_{ln\left(\frac{b_j}{\Lambda(M)k(b_j)^\alpha}\right)}^\infty \Lambda(M)A_j k_j^\alpha f(ln(A_j), M|\Omega_j)dA_j dM$$
$$- \left[1 - p\left(default_j|\Omega_j\right)\right](1+r_j)b_j - \left[p\left(default_j|\Omega_j\right)\right](1-\psi)c_d k_j \tag{1.16}$$

An increase in options ψ decreases manager's losses in case of firm's default, increasing moral hazard incentives. The firm manager internalizes the bank's supply of loan $r_j(b_j)$ and decides the optimal borrowing b_j to maximize expected payoff

$$k_j = argmaxE[w_{firm,j}(r_j(b_j), M_j, ln(A_j))|\Omega_j] \tag{1.17}$$

Appendix A.2 describes in the detail the bargaining process and the stage-2 equilibrium.

Strategic substitutability in production If $\nu < \xi$, the demand function for good j (1.1) is decreasing in aggregate production of intermediates M. For a given level of local production M_j, a lower price p_j implies lower revenues for firm j. As a result, island j's equilibrium debt b_j and loan rate r_j depend positively on the realization of local technology A_j, but negatively on the aggregate production M.

1.3.3 Stage 1: Information choice

Before observing any signal, each island decides whether to pay an information cost c to perfectly observe aggregate shock θ stage 2,which is informative about aggregate production M. Similarly to the lending decision in sage 2, I assume bank and firm managers share information and decide cooperatively through Nash bargaining with the firm holding all the bargaining power.[11] As a result, island j decides collectively to pay the attention cost if

$$E[w^*_{firm,j}(\theta \in \Omega_j, \lambda) - c] \geq E[w^*_{firm,j}(\theta \notin \Omega_j, \lambda)] \tag{1.18}$$

where w^*_{firm} are the equilibrium payoffs of firm managers in the second stage and expectation are conditional only on priors, as agents have no access to any signal at this stage. Expected profits depends on local and aggregate information choice: (i) whether managers in j will be able to observe aggregate shocks in the next stage, $\theta \in \Omega_j$, or not, $\theta \notin \Omega_j$; (ii) on the total share of islands deciding to observe aggregate shocks in the next stage $\lambda \in [0, 1]$, where $\lambda = 1$ if all islands decide to pay the cost to observes aggregate shocks and $\lambda = 0$ if none decides so. In equilibrium, λ^* is such that all island are indifferent between paying the cost or not, $E[w^*_{firm,j}(\theta \in \Omega_j, \lambda^*) - c] = E[w^*_{firm,j}(\theta \notin \Omega_j, \lambda^*)]$.

Information choice also exhibits strategic substitutability, meaning that a higher share of informed island λ decreases island j's incentive of paying the information cost. I provide intuition for this in the next section.

1.4 Credit booms and Inattention

1.4.1 Analytical results

In order to provide intuition for the model mechanism, I consider a first order approximation of the second stage model around the risky steady state (Coeurdacier

[11] Any private information between agents in the same island would be perfectly revealed by local prices. Therefore any individual decision on whether to observe private information would need to account for this information spillover, introducing strategic considerations between agents in the same island. To avoid this, I use a Nash bargaining setting where the decision is taken cooperatively with the same bargaining power as in stage-2 bargaining. As a result, the firm manager gets the surplus and pay the information cost. I allow for a different split of surplus and cost in the appendix.

et al., 2011).[12] At the steady state, all islands observes the same signal $z_j = 0$ and the aggregate shock $\theta = 0$, but there is still uncertainty about the local shock realization ϵ_j. This risk is priced in the steady state spread $r_j > r^f$, meaning there is a positive steady state risk premium. In this section I assume for simplicity no adjustment cost $\phi = 0$, no moral hazard $\psi = 0$ and no default cost $c_d = 0$. Because of these assumptions, the equilibrium perceived default risk and risk premium are constant (while actual default risk might not be), but the remaining qualitative implications of the model are unaffected. I relax all these assumptions in section 1.4.2 where I solve the full model numerically.

Proposition 1 (Linearized model) *Consider the first order approximation of the second-stage equilibrium defined by equations (1.15) and (1.17) assuming $\phi = 0$, $\psi = 0$ and $c_d = 0$. Let $\hat{x}$ indicate the log-deviation of any variable x from its steady state value and with $\tilde{x}$ the level deviation from steady state.*

- *Equilibrium local investment equals*

$$\hat{k}_j = \frac{1}{1-\alpha}(E[lnA_j|\Omega_j] - \gamma E[\hat{M}|\Omega_j]) \tag{1.19}$$

where $\hat{M} = \mu(\theta + \alpha\hat{K})$, with $\mu > 0$ and $\hat{K} = \int^j \hat{k}_j dj$. Let Ω_j denote the information set of island j (bank and firm) and $\gamma \equiv \frac{\nu-\xi}{1-(1-\alpha)\xi}$ the elasticity of the operating profit $\pi_j(A_j, k_j, M)$ with respect to aggregate production M. if $\nu < \xi$, then$\gamma < 0$ and the economy exhibits strategic substitutability in firms investment decisions. If $\nu > \xi$, then $\gamma > 0$ and the economy exhibits strategic complementarity in firms investment decisions.

[12] While the economy at the proximity of the steady state is not suitable to study large and rare financial crises like the one considered in this paper, the basic model mechanism does not rely on non-linearity and there preserve its main idea in the linearized version

- *The loan rate is proportional to perceived default risk*

$$\hat{r}_j \propto -\hat{p}(def_j|\Omega_j) \tag{1.20}$$

where $\hat{p}(def_j|\Omega_j)$ is the perceived default risk of island j conditioning on information set Ω_j.

- *Equilibrium perceived default risk is constant*

$$\hat{p}(def_j|\Omega_j) = 0 \tag{1.21}$$

- *Equilibrium aggregate bank's profits in state θ equal*

$$E[\tilde{\pi}_{bank}|z_j, \theta] \propto -\int^j [\hat{p}(def_j|z_j, \theta) - E[\hat{p}(def_j|\Omega_j)|\theta]]dj \tag{1.22}$$

where $\hat{p}(def_j|z_j, \theta)$ is the default risk conditional on signal z_j and aggregate shock θ, which I define as actual default risk.

See the appendix for the derivations.

First, notice that since I assume $\nu < \xi$, then $\gamma > 0$ and therefore equation (1.19) represents a linear game with strategic substitutability: higher aggregate investment (or debt) $\hat{K}$ lowers island j's optimal investment $\hat{k}_j$. Second, the equilibrium loan rate $\hat{r}_j$ is negatively related to the perceived default probability. This result follows directly from the price equation (1.15) and it implies that changes in risk premia only reflects changes in perceived quantity of risk. Second, perceived default risk is constant in equilibrium (or zero in log-deviation from the steady state). This is a knife-edge result that depends on the simplifying assumptions introduced in this section, which I relax in the numerical solution. Finally, as the loan pricing condition implies no expected profits for the bank, aggregate bank profits in state θ

depend on whether agents correctly perceived risk, i.e. the loan is correctly priced conditioning on θ.

PE vs GE A positive aggregate shock θ has two effects on equilibrium investment: a partial equilibrium effect and a general equilibrium effect.[13]

$$\frac{\partial \hat{k}_j}{\partial \theta} = \frac{1}{1-\alpha}\left(\underbrace{\frac{\partial E[lnA_j|\Omega_j]}{\partial \theta}}_{\text{PE effect}} - \gamma \underbrace{\frac{\partial E[\hat{M}|\Omega_j]}{\partial \theta}}_{\text{GE effect}}\right) \tag{1.23}$$

First, local productivity A_j in each island increases. Because firm's fundamental is higher, island j manager's posterior probability of default decreases, boosting borrowing and investment $\hat{k}_j$. This is the standard positive channel of productivity shocks in the existing literature and it does not depend on the interaction between islands (PE). Second, higher aggregate supply of intermediates can imply lower or higher demand for intermediate good j depending on the degree of decreasing return to scale (ν) with respect to the elasticity of substitution between intermediates (θ). Since I assume $\nu < \xi$, $\gamma > 0$ and the higher competition in the intermediate good market implies a lower demand and revenues for firm j. As a result, optimal investment $\hat{k}_j$ is lower.

While λ depends endogenously on the sage-1 information choice, I consider here two limit cases to illustrate the mechanism of the model. First, I assume all islands decide to pay attention to aggregates in the first stage ($\lambda = 1$, *full information*). Second, I assume no island decide to pay attention to aggregates in the first stage ($\lambda = 0$, *dispersed information*).

[13] Here I use partial equilibrium effect to refer to an effect related only to the island j's problem, and the term general equilibrium effect to indicate an effect related to the interaction between islands (Angeletos and Lian, 2017).

Full information $\lambda = 1$

Consider the full information case, meaning all islands decide to observe aggregate shock θ in the first stage in addition to the free signal z_j defined by equation (1.12).

Proposition 2 (Full information) *If $\Omega_j = \{z_j, \theta\}$, the solution to the linear game in proposition 1 is*

$$\hat{K}^{fi} = \frac{1-\gamma\mu}{1-\alpha+\gamma\mu\alpha}\theta \tag{1.24}$$

See the appendix for the proof.

After an aggregate shock, the improvement in local technology increases equilibrium aggregate debt and investment, but its effect is dampened by the endogenous decrease in intermediate good prices, which lower firms' optimal investment. The stronger the elasticity of intermediate price with respect to the increase in aggregate supply of intermediate $0 < \gamma < 1$, the stronger is the dampening force of the GE effect.

Corollary 1 (Actual default rate in FI) *If $\Omega_j = \{z_j, \theta\}$, actual default risk coincides with perceived default risk, which is constant by proposition 1.*

$$\hat{p}(def_j|z_j, \theta) = \hat{p}(def_j|\Omega_j) = 0 \tag{1.25}$$

As a result the default rate, which equals the average actual default risk across firms, is also constant.

Notice that the negative endogenous GE effect on expected firm's revenue can not be larger than the positive PE effect in full information, which implies that the actual default risk can not be larger either. If that was the case, then the lower expected revenues would lead the managers to decrease debt and investment (proposition 1), resulting in lower aggregate supply, higher price and a positive endogenous GE effect. In other words, if the default risk was higher, the agents in the economy would optimally limit leverage and reduce it.[14] As a result, the full information economy is not riskier in credit boom, which is at odds with the existing empirical evidence (Schularick and Taylor, 2012; Krishnamurthy and Muir, 2017).

Corollary 2 (Bank's profit in FI) *If $\Omega_j = \{z_j, \theta\}$, bank's profit are zero conditioning on z_j and θ.*

$$E[\tilde{\pi}_{bank}|z_j, \theta] = 0 \tag{1.26}$$

Because perceived risk coincides with actual risk, default risk is correctly priced conditioning on aggregate economic conditions. In other worlds, because banks observe θ, they do not make systematic errors conditioning on it. The zero expected profit condition implies that banks make zero excess return on average for each θ. While this model implies zero expected profits for banks, a different bargaining power could imply positive profits. However bank shareholders would not accept predictable losses, which is at odds with the evidence in Baron and Xiong (2017).

Dispersed information $\lambda = 0$

Consider the dispersed information case, meaning no island decides to observe aggregate shock θ in the first stage, so they only observe the free signal z_j defined by equation 1.12.

Proposition 3 (Dispersed information) *If $\Omega_j = \{z_j\}$, the solution to the linear game in proposition 1 is*

$$K^{di} = \frac{(m - \gamma\mu\delta)}{1 - \alpha + \gamma\mu\alpha\delta}\theta \tag{1.27}$$

where $m = \frac{\sigma_e^2 + \sigma_\theta^2}{\sigma_e^2 + \sigma_\theta^2 + \sigma_\eta^2}$ and $\delta = \frac{\sigma_e^2}{\sigma_e^2 + \sigma_\theta^2 + \sigma_\eta^2}$ are the Bayesian weights on signal z_j in the posterior means of $ln(A_j)$ and θ respectively, with $0 < \delta < m < 1$.

[14] This is a consequence of the strategic substitutability game between firms. A large body of research focuses instead on strategic complementarity to rationalize the procyclical leverage in full information (Gertler and Kiyotaki, 2010; Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2019; Bianchi and Mendoza, 2020).

See the appendix for the proof.

Corollary 3 (Boom amplification) *The difference in aggregate investment in dispersed information 1.27 and full information 1.24 depends positively on θ, and therefore the information friction leads to an amplification of credit booms if*

$$(m - \gamma\mu\delta)(1 - \alpha + \gamma\mu\alpha) > (1 - \gamma\mu)(1 - \alpha + \gamma\mu\alpha\delta) \tag{1.28}$$

The aggregate shock θ affect both the local fundamental (PE effect) and the aggregate production (GE effect). As a result, not observing θ leads to an underestimation of both, with opposite effects on optimal investment. Whether investment in dispersed information is larger than in full information depends on how much observing aggregates increases (i) posterior belief on local productivity (PE) and (ii) posterior belief on aggregate production of intermediates (GE).

First, suppose the signal z_j is infinitely noisy, $\sigma_\eta \to \infty$, then $m = \delta = 0$ and the condition doesn't hold. Without signals on local productivity, the aggregate shock is the only source of information. If agents do not observe it either, investment equals the steady state level in every states. If agents are instead able to observe it, higher aggregate shock θ increases their posterior on both local technology (PE) and aggregate investment (GE), but the only equilibrium is one in which the first prevails on the second and optimal local investment increases.[15] Second, suppose the signal z_j is noiseless, $\sigma_\eta \to 0$, then $m = 1, \delta < 1$ and the condition holds. In this case agents observe perfectly local productivity regardless of their information on aggregate shock. However, observing aggregates is informative on the investment decisions of the other firms in the economy, and therefore on the negative endogenous GE effect. In the dispersed information setting, after an aggregate shock

[15] To see it, suppose that the negative GE force from higher aggregate investmentt was stronger than the positive PE effect from higher local technology an optimal local investment decreased in θ. Aggregate investment would then be inversely related to θ, and the GE force would be positive for the island and not negative, leading to a contradiction.

agents underestimate the increase in competition and over-invest with respect to the economy with informed agents. This result suggests that information frictions can lead to amplification by dampening negative GE effect, similarly to Angeletos and Lian (2017).

Now consider the case of an individual island, both bank and firm, forming expectation on local firm's operating profits. Define the forecast errors as the difference between realized and expected revenues, $fe \equiv \widehat{\pi}(A_j, k_j, M) - E[\widehat{\pi}(A_j, k_j, M)|\Omega_j]$.

Corollary 4 (Rationally extrapolative beliefs and underreaction) *If $\Omega_j = \{z_j\}$, the average forecast errors on firm's revenues in state θ is proportional to*

$$E[\widehat{\pi}_j|z_j,\theta] - E[E[\widehat{\pi}_j|z_j]|\theta] =\propto -[(m-\gamma\mu\delta)(1-\alpha+\gamma\mu\alpha) - (1-\gamma\mu)(1-\alpha+\gamma\mu\alpha\delta)]\theta \tag{1.29}$$

while the forecast error on aggregate output is

$$E[\widehat{Y}|z_j,\theta] - E[E[\widehat{Y}|z_j]|\theta] = (1-\gamma\mu)\left(\frac{1-\alpha+\alpha m}{1-\alpha+\gamma\mu\alpha\delta}\right)\theta \tag{1.30}$$

If condition(1.28) holds, then

- $\theta > 0$*: agents underestimate aggregate output and overestimate individual revenues (overoptimism).*
- $\theta < 0$*: agents overestimate aggregate output and underestimate individual revenues (overpessimism).*

Equilibrium revenues depend positively on the PE effect and negatively on the GE effect. Because agents do not observe aggregates, they rationally confound an aggregate shock for a local shock and underestimate the negative GE effect. The information incompleteness produces extrapolative-like beliefs, as agents are systematically overoptimistic after a positive shock and overpessimistic after a negative shock. Differently from behavioral models where overoptimism originates from overreaction to positive news (Bordalo et al., 2018b, 2019), here it is due to rational underreaction to the endogenous negative general equilibrium effect. As a result, in booms we observe both overoptimism about local revenues and underestimation of aggregate quantities, consistently with the evidence in section 1.2. Even if agents are rational and correct on average conditioning on their information set, they can be consistently mistaken conditioning on unobserved aggregate states.

Figure 1.4 illustrates the intuition. The dotted line represents the prior belief about firm's revenues before receiving any information. A positive technology shock increases firm's fundamentals and implies on average a good signal z_j that shifts the posterior beliefs on revenues to the blue solid line (positive PE effect). However, because of the endogenous increase in intermediate good supply, price of good j will be lower and the actual posterior revenues of an informed agent would shift back to the middle dashed line (negative GE effect). Inattentive agents underestimate left tail risk, illustrated in the figure as the shaded area between their posterior and the actual posterior distribution of revenues.[16]

Corollary 5 (Actual default rate in DI) *If $\Omega_j = \{z_j\}$, the equilibrium default rate is proportional to*

$$\hat{p}(def|\theta) \propto [(m - \gamma\mu\delta)(1 - \alpha + \gamma\mu\alpha) - (1 - \gamma\mu)(1 - \alpha + \gamma\mu\alpha\delta)]\theta \tag{1.31}$$

where $\hat{p}(def|\theta) = \int^j \hat{p}(def_j|z_j, \theta)dj$. If condition (1.28) holds, default rate increases in aggregate shock θ

See the appendix for the proof.

As dispersed information amplifies the credit boom, the larger supply of intermediates lowers further prices and firms' revenues. As agents confound the aggregate

[16] Notice that, because more information also implies lower posterior uncertainty, the difference between informed and non-informed posterior is not only lower posterior mean but also lower posterior variance.

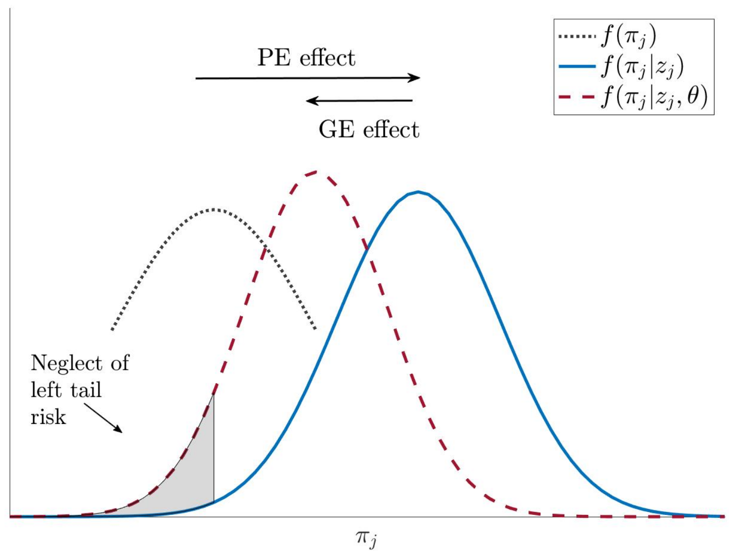


Figure 1.4: Rationally extrapolative beliefs in booms

Notes: The figure illustrates the posterior belief on firm's operating profits after a positive aggregate shock under three different information sets. The black dotted line represents the posterior of an agent not observing any new information. The blue solid line represents the posterior of an agent obsereving only local signal z_j. The red dashed line represents the posterior of an agent observing both local signal z_j and aggregate shock θ. Not observing aggregate shock θ leads to overestimating equilibrium price p_j and therefore individual revenues π_j.

for a local shock, they do not internalize this risk and increase leverage too much with respect to their repayment capacity, leading to higher default rate. As a result, credit booms are period in which default risk is larger, consistent with the evidence that low risk premium and high credit growth predict higher financial fragility (Krishnamurthy and Muir, 2017).

Corollary 6 (Bank's profit in DI) *If $\Omega_j = \{z_j\}$, the equilibrium average bank profits are proportional to*

$$E[\tilde{\pi}_{bank}|z_j, \theta] \propto [(m - \gamma\mu\delta)(1 - \alpha + \gamma\mu\alpha) - (1 - \gamma\mu)(1 - \alpha + \gamma\mu\alpha\delta)]\theta \qquad (1.32)$$

If condition (1.28) holds, average bank profits are negative after a credit boom.

Because in equilibrium the risk premium is such that banks get zero expected profit on average, when banks underestimate default risk they misprice loans and get negative profits.[17] This result is consistent with the evidence that credit booms predict negative returns on bank stocks in Baron and Xiong (2017).

Information choice In the first stage, managers decide whether they want to observe aggregates based on their expected profits in the final stage. In general, a share $\lambda \in [0,1]$ of islands decides to acquire the information. While figure 1.4 illustrates individual beliefs for a given aggregate production M, this quantity is endogenous to the aggregate amount of information in the economy. If all managers in each island are informed, $\lambda = 1$, proposition 3 states that the increase in aggregate supply in boom is lower, and the decrease in price as well. In figure 1.4, this would mean a shorter distance between informed and uninformed posterior, as the neglected GE effect is lower. On the other hand, if managers and firms in each island are uninformed, $\lambda = 0$, the credit boom is amplified, and the decline in price is larger. In figure 1.4, this would mean a larger distance between informed and uninformed posterior, as the neglected GE effect is higher. Therefore, the benefit of information for the individual island depends negatively on the average level of information in the economy. In particular, there is strategic substitutability in information choice, as higher aggregate information implies lower individual benefit of information. In section 1.5 I illustrate numerically how moral hazard incentives also affect benefit of information and equilibrium λ.

1.4.2 Numerical illustrations

I provide a numerical illustration of the non-linear model. The contribution of studying numerical solutions of the model is twofold. First, I relax some parametric assumptions needed to keep the analytical model tractable. Second, non-linear global solution are more suitable than approximation around the steady state to

[17] Since I assumed $\alpha = 0$, the condition applies to both shareholders and managers. If $\alpha > 0$, then the pricing condition is on manager's profits.

analyze the nature of large and rare credit booms, as the ones considered in this paper.

Calibration Table 1.1 reports the model's calibration. First, I set $\xi = 0.833$ to match a markup of 20%, which is inside the set of values estimated in the macro literature (for a review, see Basu (2019)). Together with a capital share $\tilde{\alpha} = 0.33$, it implies $\alpha = \frac{\tilde{\alpha}\xi}{1-(1-\tilde{\alpha})\xi} = 0.624$. The return to scale of final good producer ν can be expressed similarly as a function of the final good sector markup and the intermediate good share in production. Assuming the latter equal 0.5 (approximately the average value for the US economy over a long period of time) and a markup of 50% gives $\nu = 0.5$. The larger markup in the retail and wholesale sectors with respect to other sectors is in line with the evidence in De Loecker et al. (2020). However, my modeling assumption of $\nu < \xi$ would be satisfied by any final good sector markup larger than 13%.[18]

Since TFP in my model is i.i.d., I set the aggregate volatility equal to the unconditional volatility implied by a standard autoregressive process with quarterly shock volatility 0.02 and autoregressive coefficient 0.995, which gives $\sigma_\theta = 0.2$. I set the idiosyncratic TFP volatility $\sigma_e = 3\sigma_\theta$, where the ratio 3 is somewhere between the macro structural estimates (e.g. ≈ 15, Maćkowiak and Wiederholt (2015)) and the micro empirical estimates (e.g. ≈ 1.1, Castro et al. (2015)). Moreover, I set the private noise $\sigma_\eta = \sigma_a$, where σ_a is the total volatility of TFP. Because the model aims to capture low frequency credit boom&busts as in the macro-finance empirical literature, I set the risk free rate to the 5-year implied return from a one-year T-bill of 2%, which gives $r^f = 0.1$. The corporate tax rate is set to 20% (CBO, 2017).

In this section I abstract from manager convex compensation incentives and

[18] Assume the final good sector face a demand given by $P = Y^{\tilde{\xi}-1}$ and have a production function $Y = M^{\tilde{\nu}}X^{1-\tilde{\nu}}$, where X is some other variable input. After maximizing X out, the profit function would be proportional to $\pi \propto M^{\frac{\tilde{\nu}\tilde{\xi}}{1-(1-\tilde{\nu})\tilde{\xi}}} \equiv M^{\nu}$. Given an intermediate share of $\tilde{\nu} = 0.5$ and $\xi = 0.833$, the condition $\nu < \xi$ implies a final good sector markup $\frac{1}{\tilde{\xi}} > 1.13$.

Table 1.1: Calibration

Parameter	Interpretation	Value
α	Return to scale intermediate good sector	0.624
ν	Return to scale final good sector	0.5
r^f	Risk free rate	0.1
ϕ	Investment adj cost coefficient	1
σ_θ	Volatility aggregate shock	0.2
σ_e	Volatility local shock	0.6
σ_η	Volatility signal noise	0.64
ψ	Compensation convexity	0
c_d	Default cost	0.5
τ	Corporate tax	0.20
c	Information cost	0.0017

set $\psi = 0$. In section 1.5 I increase convex compensation incentives and study its implications on lenidng and information choice. Finally, I calibrate the cost of information c such that with no convex compensation incentives it is optimal for all islands to be collect information ($\lambda = 1$), which corresponds to around 3% of firm's dividends in the full information economy.

Full information $\lambda = 1$ Consider the full information case, where all islands decide to observe aggregate shock θ in the first stage. The blue dashed lines in figure 1.5 reports the response of aggregate credit $B = \int^j b_j dj$ (proportional to aggregate investment), average risk premium $R - r^f$, default rate and average bank profits in this economy as functions of standard deviations of the aggregate shock θ. The figure confirms the analytical results in the previous section, as large values of aggregate shock θ are associated with a credit boom. Differently from the linear model in the previous section, I allow for a non-zero investment adjustment cost. As a consequence, the probability of default is not constant but declines after boom and, because agents know the risk is lower, the risk premium declines as well. Risk is correctly priced and banks make zero average profits in boom-and-busts. The model's implications are qualitatively similar to a benchmark model that abstracts from strategic interactions between firms, but with the price externality dampening the boom.

The model is not consistent with the existing evidence. First, Schularick and Taylor (2012) show that booms are periods where financial risks accumulates, which in my model would imply a larger default rates after a credit boom. Second, Krishnamurthy and Muir (2017) document that low risk premium predict financial crisis, but in full information, because risk is correctly priced conditioning on aggregates, risk premia are positively correlated with default risk. Finally, Baron and Xiong (2017) document that average excess return on bank stocks is negative after a boom, while informed bank in the model would not accept to make average negative returns.[19]

Dispersed information $\lambda = 0$ Consider the dispersed information case, where no island decides to observe aggregate shock θ in the first stage. The red solid lines in figure 1.5 reports the response of aggregate credit $B = \int^j b_j dj$, average risk premium $R - r^f$, default rate and average bank profits in this economy as functions of standard deviations of the aggregate shock θ. The figure confirms the analytical results in the previous section. Because agents are unaware of the negative GE effect, the credit boom is amplified, as depicted by the solid red line in the upper left panel . The excess supply of intermediate goods lowers the price and revenues, but firms are inattentive to aggregates and take on too much debt. Default risk now peaks after credit booms, consistent with the evidence on credit boom-and-busts

[19] While it would be possible to set up a model where firms had higher risk tolerance and were willing to take on more risk during credit booms, bond pricing equation (1.15) implies that the risk premium would increase as a consequence, inconsistently with the evidence in Krishnamurthy and Muir (2017). If the bankers had higher risk tolerance in booms as well, risk premia could be lower in periods of high risk (e.g. Krishnamurthy and Li 2021), but it would still not be possible to have rational bankers accepting negative excess returns on average, as documented in Baron and Xiong (2017).

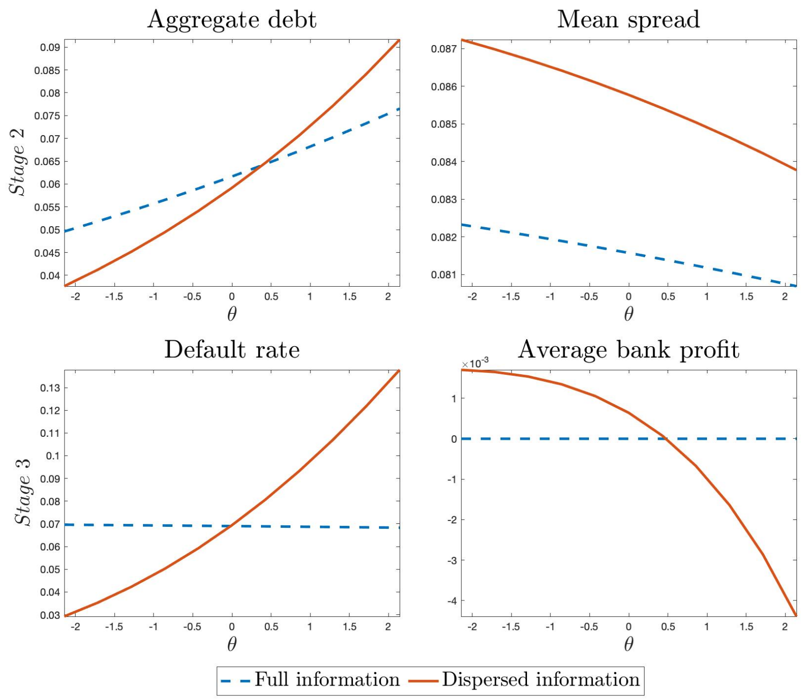


Figure 1.5: Model: full and dispersed information

Notes: The figure illustrates the equilibrium of stage-2 investment and borrowing choice in the full information ($\theta \in \Omega_j$) and dispersed information economy ($\theta \notin \Omega_j$). The aggregate shock θ in the x-axis is expressed in standard deviations.

(Schularick and Taylor, 2012). Banks are also inattentive to aggregates and they confound the aggregate shock for a local shock. As a result, the risk premium on lending is lower in credit booms when the default risk is larger. The model's results are consistent with existing evidence that high credit and low risk premia predict subsequent financial downturn (Krishnamurthy and Muir, 2017).

The decline in risk premium is not due to a change in risk tolerance, but to the underestimation of the endogenous increase in default risk. Figure 1.6 clarifies this point by plotting actual bank's profits (solid red) and mean bank's expected profits (dotted blue) in the left panel and actual average default rate (solid red) and mean expected default rate (dotted blue) on the right. Managers do not internalize the increase in default risk and expect zero average excess return. However, because of the increase in default risk, excess returns during credit booms are negative on

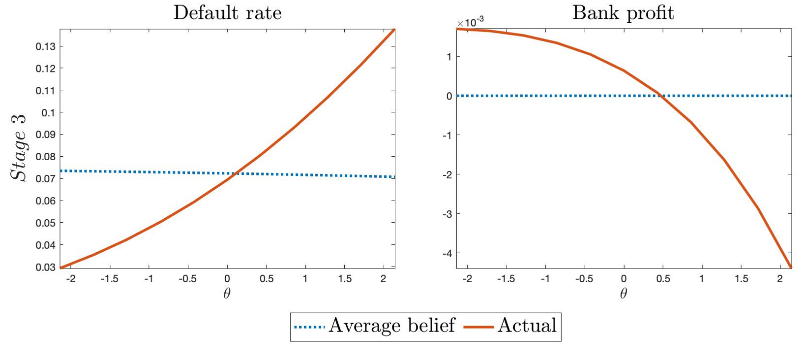


Figure 1.6: Model: actuals and beliefs

Notes: The figure illustrates the actual and the average expectation of bank excess return and default rate in the dispersed information economy ($\theta \notin \Omega_j$). The aggregate shock θ in the x-axis is expressed in standard deviations.

average. Under the assumption that bank's stock price is correlated with operating profits, the results is in line with the evidence of average negative returns on bank's stock during booms in Baron and Xiong (2017).

The equilibrium share of informed islands λ is endogenous and depends on the optimal attention decisions in stage 1. While it would be possible to rationalize a high level of information friction with a high enough information cost c, such high cost might not be realistic. I the next section I show how moral hazard incentives lead to lower optimal attention choices and therefore explain a high level of information dispersion in the model even when information costs are low.

1.5 Inattention and moral hazard

While the previous section illustrates how information frictions explain the observed frothiness and overoptimism in credit booms, I now turn to the determinant of such information friction. I show that managers with moral hazard incentives optimally decide to be inattentive to aggregates even for low information costs, causing them to be overoptimistic in booms and overpessimistic in busts. I connect the moral hazard narratives of the excessive risk taking before the financial crisis of 2008-2009 (e.g. Blinder 2009) with the behavioral overoptimism view (e.g., Gennaioli and Shleifer 2018) by showing that overoptimism is in fact a consequence of moral hazard.

Stage 2: Moral hazard in lending An increase in compensation convexity has a standard moral hazard incentive channel on stage-2 borrowing and lending decisions. First, consider the firm manager's decisions. For a given interest rate schedule $r_j(b_j)$, the firm faces a trade-off in their debt issuance b_j between higher expected profits in the no-default states and higher default probability (equation (1.17)). Higher compensation convexity ψ lowers firm managers' losses in case of default, encouraging them to take on more risk. Second, consider bank managers' decisions. Higher compensation convexity ψ implies lower losses in case of default and therefore lower elasticity of credit spread $\frac{1+r^i}{1+r^f}$ with respect to default risk (equation (1.15)).

In order to isolate the effect of moral hazard on borrowing decisions, I initially shut down the information choice in stage 1. Figure 1.7 reports the equilibrium debt, average spread, default rate and bank's profits in an economy in full information for different values of compensation asymmetry ψ. Higher moral hazard incentives lead to higher risk taking by firm managers and lower price of risk by bank managers, resulting in higher unconditional default rate. However, similarly to the full information model in the previous section, credit booms are period where the economy is safer and default rate decreases, which is not consitent with the empirical evidence (Schularick and Taylor, 2012; Krishnamurthy and Muir, 2017). Therefore the full information model with only moral hazard incentives in stage-2 borrowing decisions is not able to match qualitatively the empirical evidence on credit cycles.

Stage 1: Moral hazard in information In the first stage, bank and firm managers in each island decide whether to pay or not the information cost to observe aggregate shocks in stage 2. Both agents benefit from information, as neglecting

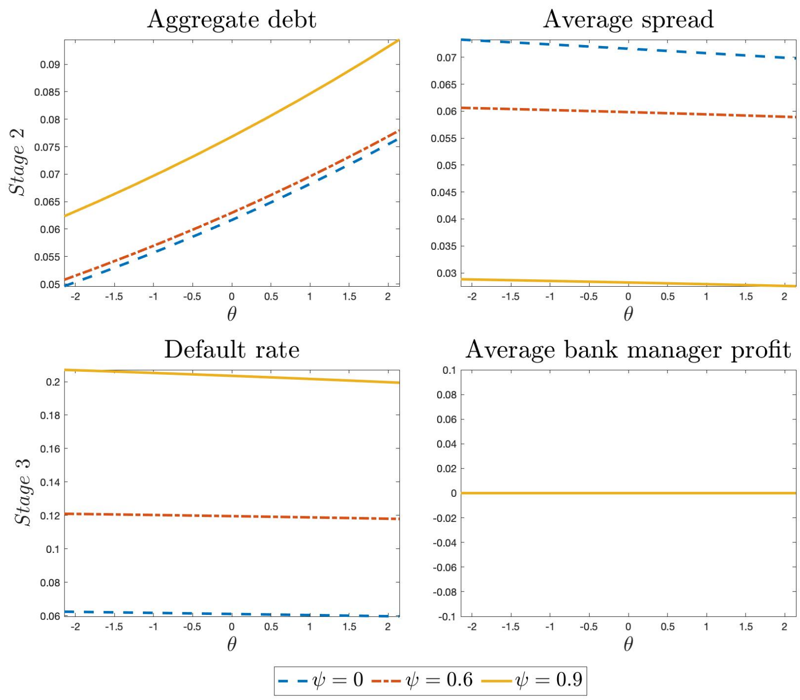


Figure 1.7: Full information and Moral hazard

Notes: The figure illustrates the stage-2 investment and borrowing choice in full information economy ($\theta \in \Omega_j$) for different values of the firm manager's compensation convexity parameter ψ. The aggregate shock θ in the x-axis is expressed in standard deviations.

aggregate shocks leads to higher default risk and losses. I set the attention cost such that, with no compensation convexity $\psi = 0$, it is optimal for all islands to pay the cost and be fully informed in next stage, $\lambda = 1$. Figure 1.8 shows that the equilibrium share of informed island λ declines in compensation convexity ψ.[20] Intuitively, the larger is the manager's compensation convexity, the lower is their exposure to losses and therefore the lower is their marginal benefit of information. [21]

[20] This result relies on the contemporaneous increase both firm and bank moral hazard incentives. First, higher firm managers' compensation convexity leads to higher risk taking and lower optimal information for a given credit spreads, but lower information also results in higher uncertainty and higher average credit spreads. Because firm managers want to take on more risk, depending on the calibration they might prefer to collect more information just to decrease price of risk. However, if bank managers' compensation convexity increases as well, then price of risk declines and the island collectively is better off with lower information.

[21] The intuition behind this result is similar to Mackowiak and Wiederholt (2012), who show that limited liability reduces optimal information choice in a general setting, while Lindbeck and Weibull (2017) study optimal contracts between principal and manager in rational inattention setting.

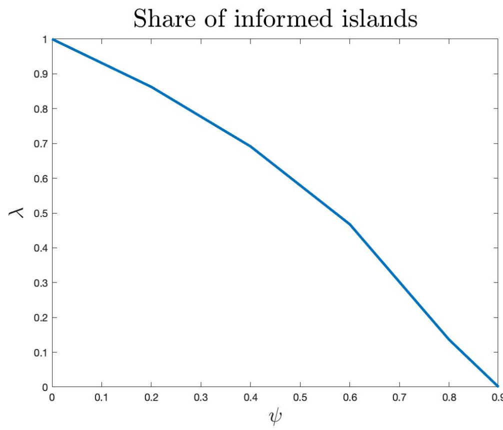


Figure 1.8: Compensation and information choice

Notes: The figure illustrates the result of stage-1 information choice under different calibration for compensation convexity ψ. It shows that higher compensation convexity is associated with lower information choice.

Figure 1.9 reports the equilibrium debt, average spread, default rate and bank's profits for different values of compensation asymmetry ψ, which endogenously lead to different value of attention λ. The higher is managers' moral hazard, the lower is the optimal attention choice, which leads to higher default rate and lower bank's profits in booms as discussed in the previous section. As a consequence, credit booms are period where default risk is larger but risk premium lower, consistently with the empirical evidence on credit cycles. The comparison between figure 1.7 and figure 1.9 reveals that moral hazard in information choice is able to explain the existing evidence on credit cycles, while moral hazard in investment decision alone is not.

1.6 Dynamic extension

In this section I extend the model to an infinite-periods setting to compare its predictions to the existing evidence on credit cycles. First, I review the existing evidence on the paths of spreads and credit before financial crises, then I compare the performance of my model against the data. While a full quantitative match of the data is beyond the scope of this paper, I show that the model is nonetheless able to produce

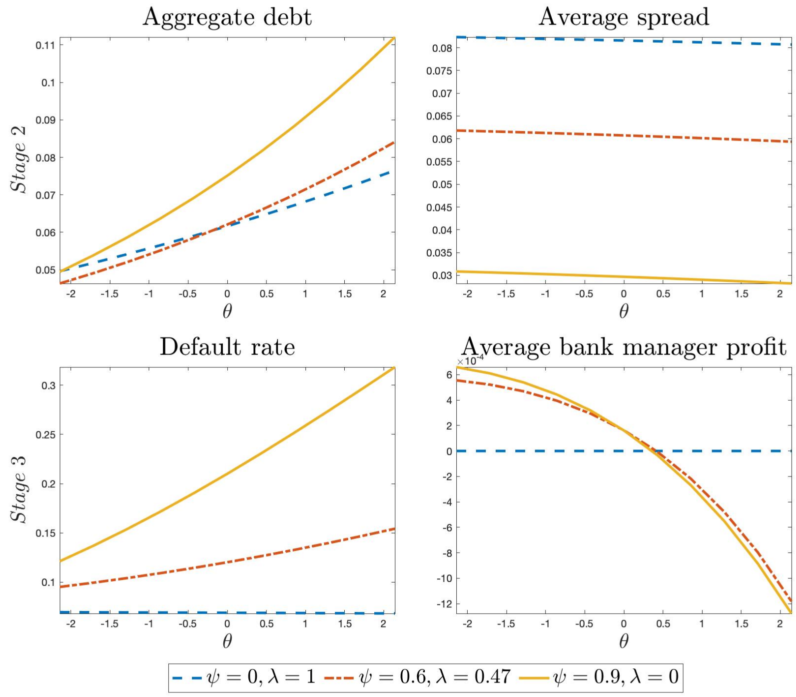


Figure 1.9: Information Choice and Moral hazard

Notes: The figure illustrates the equilibrium of the model (both stage 1 and stage 2) for different values of the firm manager's compensation convexity parameter ψ. The aggregate shock θ in the x-axis is expressed in standard deviations.

realistic boom-and-busts dynamics.

I focus on financial crises, defined by the literature "as events during which a country's banking sector experiences bank runs, sharp increases in default rates accompanied by large losses of capital that result in public intervention, bankruptcy, or forced merger of financial institutions" (Jordà et al., 2013). I compare my model against two sets of evidence from Krishnamurthy and Li (2021): first, the pre-crisis path of spreads and credit; second, the predictive power of spreads and credit growth in forecasting financial crises.

Pre-crisis period Conditioning on a crisis at time t, consider the path of spreads and credit in the 5-years preceding the crisis. First, credit spreads are 0.34σs below their country mean, where the mean is defined to exclude the crisis and the 5 years after the crisis. Second, credit/GDP is 5% above the country mean.

Predicting crises The most important evidence for the scope of this paper is the ability of spreads and credit growth to predict crises. First, Krishnamurthy and Muir (2017) find that conditioning on an episode where credit spreads are below their median value 5 years in a row, the probability of a financial crisis increase by 1.76%. Second, Schularick and Taylor (2012) shows that a one standard deviation increase in credit growth over the preceding 5 years implies an increased in probability of a crisis of 2.8% over the next year.

Dynamic model In order to related my model to the existing evidence, I embed my three-stage game in a infinite period setting. I consider an overlapping generation of bank and firm managers living for two periods. In each period a new generation of managers is born and decide information (stage 1) and lending and borrowing (stage 2). In the following period, shocks described in equation (1.7) realize, production take place and firms repay or default (stage 3). In this period, the old generation of managers receive their payoffs and die, while a new generation of managers is born and repeats the cycle.

I assume that in case of default, firms can not re-enter in the economy immediately as it takes one period for the firm to re-build its productive capacity. This simple friction can be interpreted as time needed for new firms to collect the funding to cover some fixed cost of production or to organize the production process. Define the number of defaulted firms $N_{def,t}$ as the default rate times the number of firms in the economy N_t. Then the number of firms operating in period t is given by

$$N_t = N_{t-1} - N_{def,t} + N_{def,t-1} \tag{1.33}$$

As illustrated in the previous section, in presence of moral hazard credit booms are followed by a larger default rate, which implies a lower number of productive firms in the economy in the following period. As a result, booms are followed by a burst

Table 1.2: Model and Data Moments

	Data	Model	
		$\psi = 0$	$\psi = .8$
Pre-crisis period (5 years)			
Credit spreads (σ below mean)	0.34	0.00	0.06
Credit/GDP (% above mean)	5	0	7
Predicting crises (5 years)			
Credit spreads (% increase in probability)	1.76	0.00	2.02
Credit/GDP (% increase in probability)	2.8	0.00	4.8

as in the existing evidence.[22]

In order to relate to the existing evidence on credit cycles, I calibrate one period in the model to represent a 5-years time span in the data. I follow Krishnamurthy and Li (2021) and target an annual unconditional frequency of financial crisis of 4%, which is the mean value of the different frequencies estimated in the literature. As a result, I define a financial crisis as an event in which the output drops below the 20% percentile. I solve for the model equilibrium stage-1 information and stage-2 aggregate quantities and prices for each node in 15x9 grid of aggregate shock θ_t and number of firms N_t, then I simulate 100,000 periods by drawing from the distribution of θ and interpolating from the grid. I simulate the theoretical moments for both the baseline model without compensation convexity $\psi = 0$ and with compensation convexity $\psi > 0$.

Table 1.2 reports the empirical moments and the ones generated by the model in the two different calibrations. First, the baseline model without convex compensation is not able to produce systematic movement in spreads or credit before crises,

[22] While in the framework considered here booms translates into busts through a credit demand channel, one could think of a framework where the mechanism works through a credit supply channel instead. As showed in the previous section, banks balance sheets are also impaired after booms as they suffer losses on their loans.

or to predict financial crisis with movements in spreads or credit. In this model, crises happens only when the economy is hit by negative technological shock, with no boom-and-bust dynamics. On the other hand, the model with convex compensation is qualitatively consistent with the evidence. First, crises are systematically preceded by credit boom with an increase in credit and a decline in spreads. Similarly, increase in credit and decline in spreads have predictive power on the probability of a crises in the future. Inattentive managers neglect default risk and over-invest, over-heating the economy which will end up in a recession in the following period.

1.7 Compensation and information in the data

I provide empirical support for the negative effect of compensation convexity on information choice by relating compensation of CEOs with the forecast released by their company on own earnings.

I draw mainly on three datasets. First, I collect forecast data from Institutional Brokers Estimates System (I/B/E/S) manager guidance database. This panel records in each year the forecasts released by the firm's management about their own company annual profits or earnings.[23] Second, I collect data on CEO's compensation from Execucomp and finally I get annual financial data from the Compustat/CRSP merged database. After merging these datasets, I get a panel of around 1000 CEO-firm pairs from 2004 to 2018. Appendix A.1 provides further details on the datasets and variable definition.

I define forecast errors as actual earning per share (EPS) registered by firm i in year t minus the forecasts released by firm i in year t about own EPS at the end of the same year. In order to make the errors comparable across firms, I normalize

[23] I follow the literature in measuring CEO's beliefs using firm's forecast (Otto, 2014; Hribar and Yang, 2016). The underlying assumption is that these forecasts are approved by the CEOs or, alternatively, that CEO's incentives apply also to his subordinates.

them by the standard deviation of the firm's detrended EPS.[24]

$$fe_{i,t} = \frac{eps_{i,t} - E_t[eps_{i,t}]}{sd_i} \tag{1.34}$$

I test the model's implication by regressing manager's squared forecast errors on compensation convexity.

$$\begin{aligned} fe_{i,t}^2 = \beta_0 &+ \beta_1 ln(Options_{i,t-1}) + \beta_2 ln(Shares_{i,t-1}) + \beta_3 ln(Salary_{i,t-1}) \\ &+ \beta_4 CEOcontrols_{i,t} + \beta_5 FirmControls_{i,t-1} + \eta_i + \gamma_t + \epsilon_{i,t} \end{aligned} \tag{1.35}$$

I measure compensation convexity as number of vested stock options for a given stock of equity shares and fixed salary. I use end of previous period stocks as they are more relevant for forecasts released at beginning of current period, and to minimize concerns about reverse causality. I control for CEO's age, tenure, forecast horizon and forecast width.[25] Moreover, I control for standard lagged firm financial variables.[26] Finally, I include time and CEO fixed effects.

The estimate of interest $\hat{\beta}_1$ represents the impact of an increase in stock options holding on squared forecast errors. In accordance to the model, I find a robust and statistically positive coefficient under different measures of option holdings. Table 1.3 reports the estimated $\hat{\beta}_1$ with different specifications. Column 1 uses the baseline specification, where I measure CEO's option holdings with a dummy equal to 1 if the CEO has a positive holding of stock options. Using a dummy lowers concerns about measurement errors on manager's compensation. However, In columns

[24] In order to get rid of the common trend in firm's EPS, I first subtract from each firm actual EPS the median of all other firms in the panel. Then I compute the standard deviation of the firm's EPS in on the available observations after 1985, considering only firms for which I have 10 or more years of data.

[25] As firms release forecast at different distances from the fiscal year ending date, I control for the horizon forecasted. Moreover, since some firms provide an interval and not a point forecast, I control for the forecast width.

[26] I control for annual stock return, standard deviation of returns, total assets, market capitalization, book value, leverage, stock price, total assets.

2 and 3 I use respectively value of shares and the number of shares underlying the option contracts. In column 4 I saturate the model by including 2-digits industry times year fixed effect. In all the specifications, option holding is positively and significantly correlated to squared forecast errors. This finding support the model's implications, as larger compensation convexity leads the manager to take on more risk and neglect information.

Table 1.3: Option compensation and squared forecast errors

	(1)	(2)	(3)	(4)	(5)
	fe^2	fe^2	fe^2	fe^2	fe^2
	Coef./SE	Coef./SE	Coef./SE	Coef./SE	Coef./SE
OptionsDummy	0.134**			0.107**	0.038*
	(0.055)			(0.043)	(0.019)
lnOptionsVal		0.017**			
		(0.007)			
lnOptionsNum			0.028*		
			(0.013)		
R-squared	0.062	0.062	0.063	0.089	0.824
N	4482	4475	4475	4244	4455
Controls	Y	Y	Y	Y	Y
Year fe	Y	Y	Y	Y	Y
Ceo-firm fe	Y	Y	Y	Y	Y
Year×industry fe				Y	

Note: the table reports the estimated $\hat{\beta}_1$ from regression 1.35. Column 1 reports the baseline model, where I measure options simply with a dummy having value 1 if vested option holding is positive. Column 2 measures options holding as the value of the stock underlying the option contracts. Column 3 measures options holding as the number of stocks underlying the option contracts. Column 4 includes analyst squared forecast errors as control. Column 5 includes 2-digit sector times year fixed effect. Additional controls include: CEO's characteristics, as lagged number equity shares (value in column 2), lagged fixed salary, age, tenure, forecast horizon and forecast width; lagged firm's financial variable, as stock annual return, standard deviation, market capitalization, book value, leverage, stock price (except column 2), total assets, EPS.

While the theory implies that compensation convexity increases both risk taking and inattention to risk, I isolate the latter by controlling for the squared forecast error of analysts' mean forecast. Intuitively, since manager's option compensation does not affect information choice of analysts, their forecast errors reflect only the endogenous increase in the firm's EPS volatility due to manager's risk taking, but not his information choice. Therefore, controlling for analyst forecast errors help me isolating the information channel. Column 5 in table 1.3 reports the result by including this additional control. The impact of option on manager's squared errors is lower than in column 1, as expected, but still positive and significant.[27]

Since the seminal paper of Malmendier and Tate (2005), the behavioral corporate finance literature has used manager's decision to hold vested options instead of exercising them as a measure of CEO overconfidence (or overoptimism). The CEO fixed effects in my regressions take care of any non-time varying CEO characteristics, and therefore control for CEO's intrinsic overconfidence or risk aversion.

My findings are in line with a large body of empirical works on the impact of compensation incentives on risk taking (Edmans et al., 2017) and on the impact of CEO's overoptimism on risk taking (Ho et al., 2016). The result above suggests that both CEO's beliefs and risk taking are related to compensation.

1.8 Discussion and policy implications

My model implies that inattentive agents over-accumulate debt and investment during booms, which leads to higher default risk and economic fragility. While this results is similar to a large strand of the macroeconomic literature on financial frictions, the underlying mechanism is different and it highlights novel macroprudential policy implications.

A large class of models in the macro-financial literature rationalizes the over-accumulation of debt during booms with strategic complementarity in leverage choices with full information: it is individually optimal to increase leverage when

[27] The large R squared is due to the large explanatory power than analysts squared error has on manager's squared error. Nonetheless, option holding retains some explanatory power, which is due to the information effect alone.

other agents do it, as individuals do not internalize the impact of their decision on the aggregate economy (Gertler and Kiyotaki, 2010; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2019; Bianchi and Mendoza, 2020). However, it is socially suboptimal, as it leads to high levels of leverage and financial fragility. In this framework, a Pigouvian tax on investment corrects this externality by mitigating the increase in leverage (Jeanne and Korinek, 2019).

In my model, the socially suboptimal high borrowing and investment during booms results from the combination of strategic substitutability and imperfect information. As aggregate investment increases, informed firms and banks would decrease their own lending and investment, making the economy safer. However, because they can not perfectly observe aggregates, they contribute in making the economy riskier by increasing their own lending and investment. Information provision would then mitigate the overoptimism and therefore the boom-and-busts cycles.

The policy maker can solve the information friction and mitigate the boom-and-bust cycles by correcting managers' moral hazard. While the policy maker could provide free information though public announcements or direct communication with managers, they still have to pay a cognitive cost to process this information (Sims, 2003; Mackowiak et al., 2018). However, the policy maker can affect managers' incentives to collect information by making them accountable for their mistakes in belief formation. A feasible policy in this direction is regulating managers' compensation structure by limiting stock options compensation. An example of this policy is the Tax Cuts and Jobs Act (TCJA), that in 2017 reduced the scope of tax deductability for performance-based compensation as stock options (Durrant et al., 2020).

1.9 Conclusions

I presented a theoretical framework where overoptimism originates from moral hazard incentives in information choice. While existing models explain overoptimism during credit booms with behavioral extrapolation to good news, I propose a rational framework where overoptimism originates instead from inattention to negative news. In particular, large credit booms are associated with an increase in aggregate supply and decrease in price, and therefore inattention to aggregates leads to overestimation of own revenues. As a result, managers over-borrow and over-invest, overheating further the economy. Periods of low risk premium predict higher default rate and systematic bank losses, in line with existing evidence. Moreover, I show that such information friction can result from moral hazard incentives, as convex compensation structures discourage managers to collect information. Finally, I document a positive relation between CEO's compensation incentives and information in a large sample of US firms. Because beliefs depend on incentives, my model suggests that compensation regulation has important implication in terms of macro-prudential policy.

Chapter 2

Biased Surveys

2.1 Introduction

Expectations play a crucial role in macroeconomic models, and hence the process through which agents form their expectation has been a fundamental, and often debated, topic. An important new development in the literature has emphasized the use of survey data, which holds the promise of providing direct, micro-level measurement of agent expectations. Using such data, Coibion and Gorodnichenko (2012, 2015) find significant evidence of incomplete and imperfect information, while another set of studies documents extensive predictability in individual forecast errors, which calls into question the classic paradigm of rational expectations itself (e.g. Bordalo et al. (2020)). Both strands of the literature, however, rely on the strong assumption that the information set of agents are contaminated with purely idiosyncratic errors, excluding any correlation in the noise of agent beliefs.

We empirically investigate the importance of public information, which introduces a common error component, and provide new evidence on the full information and rational expectations (FIRE) hypotheses. Our key findings are two-fold. First, we indeed find significant evidence of a common noise component in expectations, which biases the standard estimates of informational rigidity downwards. In particular, our findings indicate that prior studies have under-estimated the degree of information rigidity by about a third on average across a variety of macroeconomic indicators, and by up to 50% in the case of long-term interest rates.

Second, we find that, while individual forecasters tend to *over-react* to new information on average (in-line with previous findings of Bordalo et al. (2020)), the forecasts actually *under-react* to new public (i.e. common) information. We show that this finding is in-line with models where strategic diversification incentives lead forecasters to provide a biased measure of their actual beliefs when responding to surveys (e.g. Ottaviani and Sørensen (2006)). To quantify this effect and recover the true underlying expectations, we estimate a dynamic model of strategic incentives in reporting forecasts.

Our findings indicate that strategic incentives indeed play an important role, and hence caution against the use of survey forecasts as a direct measure of agent expectations. Specifically, the estimated model can fully account for the "over-reaction" puzzle in surveys that has received a lot of recent attention, suggesting that Rational Expectations is in fact a good model for the underlying true beliefs of agents. Moreover, the model estimates also show that the strategic incentives themselves bias the estimated information rigidity downward by a further 20% on average. Hence, our results indicate that expectations are rational after all, but the degree of imperfect information is significantly greater than previously thought.

In our empirical work we use data from the Survey of Professional Forecasters (SPF), which by now has become the common dataset for survey of macro forecasts, and we proceed in two steps. First, we show that the seminal estimate of informational rigidity of Coibion and Gorodnichenko (2015) is biased downward, due to common noise in the forecast errors. Such noise could be due to the incorporation of public signals in the forecasts, for example central bank's communications (e.g. Morse and Vissing-Jorgensen (2020)). We then provide a new empirical strategy robust to the presence of common noise by exploiting the panel dimension of the survey data. After correcting for this bias, the resulting Kalman gain estimate we find is on average 30% lower than that estimated by Coibion and Gorodnichenko (2015), revealing a significantly higher degree of information rigidity than previously found. Moreover, in the particularly interesting case of inflation expectations, we find that information rigidity is actually 40% higher than previous estimates, suggesting an even more important role for imperfect information in the transmission of monetary policy.

Second, we refine tests of rational expectations in survey data by also incorporating public signals and information in the benchmark regression specifications. In particular, we find that the *lagged* consensus forecast has an outsize importance, as it is both publicly available to all forecasters when they make their current forecasts and is also highly informative about the future realization of the variable being forecasted. We show that, while the individual forecasts in the SPF appear to over-weight new information on average (as already documented by Bordalo et al. (2020), hereafter BGMS), they significantly *under-weight* the information in the previous quarter's consensus forecast.

To rationalize our empirical findings, we build a global game model à la Morris and Shin (2002) with strategic substitutability, where the forecaster is balancing the desire to be right with the desire to stand-out. Intuitively, the forecaster would most like to both be right and also be the only person that gave a correct forecast, introducing strategic diversification incentives in forecast reporting as in Ottaviani and Sørensen (2006).[1] We assume agents have access to two types of noisy signals – a private signal with idiosyncratic noise, and a noisy public signal that is the same for everyone. We then show that, because of strategic substitutability incentives in responding to the survey, agents optimally decide to bias their response towards pri-

[1] This setting can be interpreted as a general version of a winner-take-all game, in which being accurate is rewarded but the prize is shared among correct forecasters (Ottaviani and Sørensen, 2006)

vate information, leading to overreaction to private information and underreaction to public information, as we also find in the data.

Moreover, we prove that in this setting it is always the case that individual forecasts appear as if they are *over-reacting* to new information on average, which can explain the recent findings in BGMS. Intuitively, because of agents' desire to stand out, when revising their expectations they put too high of a weight on their private signals which then results in forecastable errors that look like "over-reaction".

While BGMS ascribe this predictability of forecast errors to departures from rational expectation, we preserve rational expectation and depart instead from the assumption of honest reporting. Furthermore, in models of behavioral extrapolation agents over-react to all new information, both public and private, but this is inconsistent with our key finding that forecasts in fact significantly under-react to public information. Moreover, we further refine our test by considering variation in the cross-section of plausible public signals. In particular, in addition to the past release of the consensus forecasts, we consider another type of public information – the past realization of the macroeconomic variable being forecasted (e.g. lagged inflation). It turns out that this second type of public signal is under-weighted to a much smaller degree, which is again qualitatively consistent with the hypothesis of strategic incentives. Because the past consensus is not only a signal that ev- everyone has access to, but is also a direct estimate of everyone else's recent beliefs, strategic diversification incentives imply that it will be doubly under-weighted. Thus, we provide an alternative, rational explanation of the over-reaction evidence, that is also consistent with additional, nuanced facts we uncover.

Finally, we estimate a dynamic, quantitative version of our model which allows us to back-out and measure the actual expectations of the forecasters, after removing the estimated bias due to strategic incentives. Our key results in this section are two-fold. First, we find that the reported consensus estimate is significantly more accurate than the true average belief – with the mean-squared error of the true average belief being roughly 30% to 100% higher, depending on the variable. This result is intuitive – the simultaneous over-weighting of private information and under-weighting of public information acts as a positive externality in terms of the consensus estimate, as it limits the effects of common errors. Second, the true beliefs are also significantly less dispersed in the cross-section, with the cross-sectional standard deviation of beliefs being roughly 80% lower than the dispersion of the forecasts reported to SPF. This is also intuitive, and is a hallmark of the forecasters' attempts to "stand-out". It also speaks to the fact that the true disagreement and dispersion of beliefs is much lower than otherwise thought, and thus also consistent with an even higher degree of information stickiness.

Related literature This paper relates to three strands of the literature. First, papers using survey of professional forecasters to test the full information hypothesis. A common finding in this literature is consensus underreaction, meaning a positive relation between consensus forecast errors and consensus forecast revisions (Crowe, 2010; Coibion and Gorodnichenko, 2012, 2015). We contribute by (i) proposing an empirical strategy to consistently estimate information rigidity in presence of public information, and (ii) using the structural model to estimate the actual information rigidity of honest beliefs in presence of strategic incentives in forecast reporting.

Another strand of the literature uses surveys to test the rational expectation hypothesis. In particular, Bordalo et al. (2020) documents individual overreaction, meaning a negative relation between individual forecast errors and individual forecast revisions. As individual forecast errors should not be predictable using current information, the authors interpret this predictability as evidence of behavioral biases in belief formation. We show that this evidence can be explained by a departure from truthful revelation while preserving rational expectations. Moreover, we document underreaction to public information, which is consistent with a strategic incentive model but not with models of extrapolative beliefs. In a contemporaneous paper, Broer and Kohlhas (2018) also use public information to improve on the test of RE, and find mixed results in terms of under and over-reaction. The key difference is that in our empirical approach we isolate the surprise component of any given public signal, which leads to higher estimation precision, while they use the raw value of the public signal itself (which is correlated with other variables on the right-hand side of the main regressions).

A third group of papers investigate the role of strategic incentives in forecasters behavior (for a review, Marinovic et al. 2013). The most related is Ottaviani and Sørensen (2006), that propose two models of strategic substitutability and complementarity that leads forecasters to over or underweight private information in their reported forecast. While the spirit of the analysis is the same, we employ a more general Morris and Shin (2002) game and focus only on strategic substitutability. Moreover, we (i) introduce a dynamic model which allows use to focus on the time-series dimension of survey data, (ii) introduce public signals to distinguish between strategic incentive and behavioral theories and (iii) estimate a structural model to recover the underlying true expectations.

Overall, our results also speak to the fact that imperfect and noisy information is the dominant paradigm in the data, supporting earlier results on the importance of information rigidities in the expectation formation process, such as Kiley (2007), Klenow and Willis (2007), Korenok (2008), Dupor et al. (2010), Knotek II (2010), Coibion and Gorodnichenko (2012), and Coibion and Gorodnichenko (2015). In contrast to this literature, however, we also specifically identify and quantify the contribution of common noise components in the (imperfect) information sets of agents, and of the biasing effects of strategic incentives survey responders face when reporting expectations.

Structure of the paper The remaining sections of the paper are organized as follows. In Section 2 we describe the data and replicate the empirical findings of CG and BGMS, i.e. respectively underreaction of consensus forecast and overreaction of individual forecasts. Then, we propose a novel empirical methodology to estimate information stickiness in the presence of common noise, and also document a novel fact: forecast underreact to new public information and overreact to new private information. We show that noisy information and diagnostic expectations are not enough to explain all three facts, and therefore we turn to departure from truthful reporting. In section 3 we develop a static model of strategic substitutability in forecast reporting which can rationalize the empirical evidence and provide the additional empirical implication,. i.e. contemporaneous underreaction to new public information and overreaction to new private information. In section 4 we extend the model to a dynamic setting and estimate it, allowing us to recover honest forecast and correctly measure information rigidity.

2.2 Empirical Analysis

Data on forecast We collect data on forecasts from the Survey of Professional Forecasters (SPF), currently run by the Federal Bank Reserve of Philadelphia. In each quarter around 40 professional forecasters contribute to the SPF with forecasts for outcomes in the current quarter and the next four quarters. Individual forecasts are collected at the end of the second month of the quarter, and the forecasters are anonymous but identified by forecasters IDs.

The SPF covers macroeconomic and financial outcomes, providing both consensus forecast and an unbalanced panel of individual forecasts. These variables include GDP, price indices, consumption, investment, unemployment, government consumption, yields on government bonds and corporate bonds.

While most macroeconomic variables are provided in SPF in level, we follow

BGMS and CG in transforming them into implied growth rate. Because of the timing of the survey, the actual variable realization in $t-1$ is known to the forecasters at the time of the forecasts. Therefore, we compute the forecasted growth rate of the variable from $t-1$ to $t+3$. We apply this method for GDP, price indices, consumption, investment and government consumption, while we keep the forecast in level for unemployment and financial variables.

We winsorize outliers by removing forecasts that are more than 5 interquartile ranges away from the median of each horizon in each quarter. We keep forecasters with at least 10 observations in all analyses. Consensus forecast are computed as the average of the individual forecasts available in each quarter. Appendix B.1 provides a description of variable construction.

Data on actual outcomes The values of macroeconomics variables are released quarterly but subsequently revised. At the time of the survey, the forecasters can observe the first release of the values of the variables in $t-1$. To match as closely as possible the information set of the forecasters, we follow BGMS and use the initial releases of macroeconomics variables from Philadelphia Fed's Real-Time Data Set for Macroeconomics. Financial variables are not revised, so we use historical data from the Federal Reserve Bank of St. Louis.

We use actual realization to compute forecast errors, defined as actual realization minus forecast, and forecast revisions, defined as forecast in t on some horizon $t+h$ minus forecast in $t-1$ about the same horizon $t+h$.

Summary Statistics Table 2.1 presents the summary statistics for each series. Columns 1-5 reports the statistics for the consensus errors and revisions, including mean, standard deviation and standard errors.. Forecast errors are statistically indistinguishable from zero for most of the series except for the interest rates, for which the forecasts are systematically above realizations. As argued by BGMS, this is likely due to the downward trend of the interest rates during the sample period,

Table 2.1: Summary Statistics

	Consensus					Individual		
	Errors			Revisions		Forecast dispersion	Nonrev share	Pr(< 80% revise same direction)
	Mean	SD	SE	Mean	SD			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.26	1.69	0.19	-0.14	0.68	1.00	0.02	0.80
GDP price index inflation	-0.28	0.58	0.08	-0.08	0.25	0.49	0.07	0.85
Real GDP	-0.30	1.78	0.19	-0.16	0.58	0.78	0.02	0.74
Consumer Price Index	-0.08	1.04	0.15	-0.11	0.68	0.54	0.06	0.66
Industrial production	-0.83	3.94	0.46	-0.49	1.19	1.57	0.01	0.72
Housing Start	-3.36	17.79	2.20	-2.31	5.93	8.34	0.00	0.68
Real Consumption	0.32	1.10	0.15	-0.06	0.41	0.61	0.03	0.78
Real residential investment	-0.46	8.32	1.19	-0.61	2.33	4.37	0.04	0.87
Real nonresidential investment	0.20	5.60	0.79	-0.22	1.71	2.31	0.03	0.74
Real state and local government consumption	0.04	2.96	0.38	0.14	1.10	2.09	0.07	0.91
Real federal government consumption	0.02	1.10	0.15	-0.05	0.33	0.98	0.11	0.93
Unemployment rate	0.01	0.68	0.08	0.05	0.32	0.30	0.18	0.66
Three-month Treasury rate	-0.51	1.14	0.16	-0.19	0.51	0.43	0.15	0.59
Ten-year Treasury rate	-0.48	0.73	0.11	-0.12	0.36	0.37	0.11	0.55
AAA Corporate Rate Bond	-0.46	0.82	0.11	-0.11	0.38	0.49	0.09	0.66

Notes: Columns 1 to 5 show statistics for consensus forecast errors and revisions. Errors are defined as actuals minus forecasts, where actuals are the realized outcome corresponding to the variable forecasted. Revisions are forecast provided in t minus forecasts provided in $t-1$ about the same horizon. Columns 6 to 8 show statistics for individual forecasts, with Newey West (1994) standard errors. Forecast dispersion is the average standard deviation of individual forecasts at each quarter. The share of nonrevisions is the average quarterly share of instances in which forecast revision is less than 0.01 percentage points. The final column shows the fraction of quarters where less than 80 percent of the forecasters revise in the same direction.

to which the forecast adjust only partially.

Columns 6-8 reports the summary statistics of the individual forecasts, including forecasts dispersion, share forecast with no meaningful revisions[2] and the probability that less than 80 percent of forecasters revise in the same direction. The large dispersion of forecasts and revisions at each point in time suggest a role for dispersed information among forecasts, which we embed in our model. The share of non revisions is often small, contrary to a sticky-information model a la Mankiw and Reis (2002), and revisions go in different direction, suggestion a noisy information setting instead.

Theoretical framework We consider a general framework of belief updating with dispersed information. In particular, consider a variable following an autoregressive process

$$x_t = \rho x_{t-1} + u_t \tag{2.1}$$

where u_t is an i.i.d. normally distributed innovation to x_t with variance ξ^{-1}. Agents cannot directly observe x_t, but instead receive a private signal and a public signal

$$\begin{aligned} g_t &= x_t + e_t \\ s_t^i &= x_t + \eta_t^i \end{aligned} \tag{2.2}$$

where η_t^i represents a normally distributed mean-zero noise with variance τ^{-1} which is i.i.d. across time and across agents, while e_t represents a normally distributed mean-zero noise with variance ν^{-1} which is i.i.d. only across time, but common across agents. Each agent generates forecast $\tilde{E}_t^i[x_{t+h}]$ at time t about the variable at

[2] We follow BGMS in categorizing a non-missing forecast as a non meaningful revision if the forecasts change by less than 0.01 percentage points.

h periods ahead according to

$$\begin{aligned}\tilde{E}^i_t[x_t] &= \tilde{E}^i_{t-1}[x_t] + G_1(g_t - \tilde{E}^i_{t-1}[x_t]) + G_2(s^i_t - \tilde{E}^i_{t-1}[x_t]) \\ \tilde{E}^i_t[x_{t+h}] &= \rho^h \tilde{E}^i_t[x_t]\end{aligned} \tag{2.3}$$

where G_1 is the weight agent put on the public signal, G_2 the weight agent put on the private signal and $\tilde{E}$ is a potentially non-optimal expectation operator. This general format embeds the rational Bayesian case, where $\tilde{E}_t[x_t] = E_t[x_t]$, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with $\Sigma \equiv var(x_t - E^i_t[x_t])$.

2.2.1 Fact 1: under-reaction in consensus forecasts

Denote with $\bar{\tilde{E}}_t[x_{t+h}]$ the mean forecast across forecasters. Consensus forecast error is defined as the actual realization minus the average forecast: $\bar{fe}_{t+h,t} = x_{t+h} - \bar{\tilde{E}}_t[x_{t+h}]$. Similarly, consensus forecast revision is defined as the average forecast provided today minus the forecast provided in the previous period about the same horizon: $\bar{fr}_{t+h,t} = \bar{\tilde{E}}_t[x_{t+h}] - \bar{\tilde{E}}_{t-1}[x_{t+h}]$.

Coibion and Gorodnichenko (2015), hereafter CG, test the full information rational expectation hypothesis by regressing consensus forecast errors on consensus forecast revisions. Intuitively if information was complete and forecasters fully rational, it should not be possible to predict future errors using today's revisions, which would be in the forecasters' information sets. They run the following regression

$$\bar{fe}_{t+h,t} = \alpha + \beta_{CG}\bar{fr}_{t+h,t} + err_t \tag{2.4}$$

A positive $\beta_{CG} > 0$ would instead imply that a upward revisions today are associated on average with forecast not optimistic enough, and therefore a systematic undershooting of the actual realization of x. On the other hand, a negative $\beta_{CG} < 0$ would imply that a upward revisions today are associated on average with forecast too optimistic, and therefore a systematic overshooting of the actual realization of x. CG document that $\beta_{CG} > 0$ is a robust finding for inflation expectation, while BGMS replicate their analysis for a wide range of macroeconomic and financial series and confirm the same result. We replicate their results in Table 2.2, for both three and two quarters horizons, in our data set.

In order to interpret the result, derive the structural equivalent from 2.3:

$$\bar{f}e_{t+h,t} = \frac{1-G}{G}\bar{f}r_{t+h,t} - \frac{G_1}{G}\rho^h e_t + \varepsilon_{t+h,t} \tag{2.5}$$

where $\varepsilon_{t+h,t} = \sum_{i=1}^{h} \rho^{h-i} u_{t+i}$.

First, consider a setting without public information: $G_1 = 0$. In this case, $\beta_{CG} = \frac{1-G}{G}$. A $\beta = 0$ would imply $G = 1$, meaning forecast adjust completely to new information, as implied by the FIRE hypothesis. On the other hand, $\beta > 0$ would imply $G < 1$, meaning stickiness in forecast updating, as in the noisy information setting. Therefore, in absence of public information, $\beta > 0$ rejects full information model, but not necessarily rational expectation.

Intuitively, in a dispersed information setting individual forecasters do not observe the information of the others, and therefore the average forecast revisions can predict average forecast errors. Because of private noise the individual signal is more noisy and less accurate that the average signal, even if each individual update their forecast optimally given their signal the average forecast is suboptimally sticky with respect to the average signal.

While this intuition is accurate in absence of public information, it is not if $G_1 > 0$. Because of the bias introduce by the public noise in the regression error, β_{CG} does not identify the information gain G.

Proposition 4 *If agents forecasts follow 2.3, the coefficient from regression 2.4 is given* *by:*

$$\beta_{CG} = \frac{\tilde{\bar{\Sigma}} - [G\tilde{\bar{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}]}{G\tilde{\bar{\Sigma}} + \frac{G_1^2}{G}\nu^{-1}} \tag{2.6}$$

with $\tilde{\bar{\Sigma}} \equiv var(x_t - \tilde{\bar{E}}_{t-1}[x_t]) = \frac{\rho^2[G_1^2\nu^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$. *Ig* $G_1 = 0$, $\beta_{CG}=\frac{1-G}{G}$

Corollary 7 *If agents forecasts follow 2.3, under rational expectation the coefficient* β_{CG} *of regression 2.4 is equal to zero in either of these two cases:*

1. *Public information very imprecise and private information infinitely precise:* $\nu = 0$*,* $\tau \to \infty$
2. *Private information very imprecise:* $\tau = 0$

while it is positive in any other case.

It follows that $\beta_{CG} = \frac{1-G}{G}$ only if $G_1 = 0$. The CG regressions doesn't provide an estimate of the consensus stickiness (or gain) in the presence of public noise. A β close to zero, as in the case of Ten-year Treasury rate and AAA Corporate Rate Bond, does not imply absence of stickiness in adjustment, but either no public information with perfectly informative private information (as in CG), or no private information. We next provide an accurate measure of forecast update stickiness, which generalized CG method to public information.

2.2.2 Fact 2: stickiness with public information

We propose a different empirical strategy to recover the weight on new information G in presence of public information. Define individual forecast revision as $fr^i_{t+h,t} = \tilde{E}^i_t[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}]$ and forecast surprise as $fs^i_{t+h,t} = x_{t+h} - \tilde{E}^i_{t-1}[x_{t+h}]$. Rewrite 2.3 in terms of forecast revision on forecast errors at some horizon h

$$fr^i_{t+h,t} = G(fs^i_{t+h,t}) + G_1\rho^h e_t + G_2\rho^h\eta^i_t - G\sum_{i=1}^{h}\rho^{h-i}u_{t+i} \tag{2.7}$$

Table 2.2: Consensus errors on revisions

	3 quarters horizon			2 quarters horizon		
	β	SE	p-value	β	SE	p-value
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Nominal GDP	0.52	0.31	0.09	0.20	0.11	0.07
GDP price index inflation	0.29	0.22	0.18	-0.03	0.11	0.77
Real GDP	0.65	0.20	0.00	0.33	0.16	0.05
Consumer Price Index	0.22	0.25	0.39	-0.07	0.09	0.44
Industrial production	0.21	0.55	0.70	0.33	0.23	0.15
Housing Start	0.38	0.25	0.13	0.91	0.14	0.00
Real Consumption	0.31	0.33	0.35	0.03	0.13	0.81
Real residential investment	1.22	0.32	0.00	0.56	0.14	0.00
Real nonresidential investment	1.21	0.21	0.00	0.52	0.09	0.00
Real state and local government consumption	-0.23	0.19	0.22	-0.28	0.17	0.10
Real federal government consumption	0.63	0.33	0.06	-0.11	0.08	0.20
Unemployment rate	0.74	0.16	0.00	0.58	0.11	0.00
Three-month Treasury rate	0.62	0.17	0.00	0.41	0.12	0.00
Ten-year Treasury rate	-0.01	0.09	0.91	-0.09	0.15	0.57
AAA Corporate Rate Bond	-0.03	0.17	0.88	0.05	0.11	0.64

Notes: The table shows the coefficient of the CG regression (consensus forecast errors on consensus revisions) with standard errors and corresponding p-values. Columns (1)-(3) consider a forecast horizon of 3 quarters, while columns (4)-(6) consider a forecast horizon of 2 quarter. Standard errors are robust to heteroskedasticity and Newey-West with the automatic bandwidth selection procedure of Newey and West (1994).

The estimated coefficient from regressing individual forecast revisions on individual surprise does not converge to G, as it is biased by the correlation between x_{t+h} and the sum of fundamental disturbances from $t+1$ to $t+h$. However, by demeaning 2.7 at every t one gets

$$(fr^i_{t+h,t}) - (\bar{f}r_{t+h,t}) = G(\bar{\tilde{E}}_{t-1}[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}]) - G_2\rho^h\eta^i_t \tag{2.8}$$

Equation 2.8 provide an unbiased strategy to measure information stickiness. The coefficient estimated by regressing the difference between individual and consensus forecast revision on the difference between individual and consensus prior converge to the posted weight on new information G. In particular, consider the regression

$$fr^i_{t+h,t} - \bar{f}r_{t+h,t} = \beta(\bar{\hat{x}}_{t+h,t-1} - \hat{x}^i_{t+h,t-1}) + err^i_t \tag{2.9}$$

the OLS coefficient $\hat{\beta}$ is an efficient estimator of gain G. This approach is more general than the CG regression as it doesn't rely on the assumption of no public information.

We run regression 2.9 in a panel data with time fixed effect to demean forecast revisions and priors at each quarter. Table 2.3 reports the estimated coefficient, standard errors and p-value of the panel data regression, and the median coefficient from demeaned individual regressions. We estimate the gains for both 3 quarters and 2 quarters horizons. There are two important observations. First, our estimated gains are relatively stable across variables at the same horizon. Second, the gains are systematically larger at the shorter horizon, consistently with the idea of more accurate information about shorter horizons.

In absence of public information, the gain $G_{CG} = \frac{1}{1+\beta_{CG}}$ implied by the CG regression 2.4 should equal the gain estimated directly from regression 2.9. However, in presence of public information, the former is larger than the latter.

Proposition 5 *If agents forecasts follow 2.3, the difference between $G_{CG} \equiv \frac{1}{1+\beta_{CG}}$, where β_{CG} is the coefficient from regression 2.4, and G, where G is the coefficient from regression 2.9, is given by:*

$$G_{CG} - G = G\left(\frac{G_1^2\nu^{-1}}{G^2\bar{\tilde{\Sigma}}}\right) > 0 \tag{2.10}$$

with $\bar{\tilde{\Sigma}} \equiv var(x_t - \tilde{E}_{t-1}[x_t]) = \frac{\rho^2[G_1^2\nu^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$.

The two estimated gains are the same if $G_1^2 = 0$, meaning no public information in agents forecast updating. However, if agents have access to public information,

Table 2.3: Stickiness estimation

	3 quarters horizon				2 quarters horizon			
	β	SE	p-value	Median	β	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	0.53	0.02	0.00	0.49	0.61	0.01	0.00	0.62
GDP price index inflation	0.49	0.03	0.00	0.52	0.63	0.02	0.00	0.68
Real GDP	0.56	0.03	0.00	0.52	0.63	0.02	0.00	0.62
Consumer Price Index	0.49	0.02	0.00	0.53	0.70	0.02	0.00	0.71
Industrial production	0.50	0.03	0.00	0.52	0.59	0.02	0.00	0.63
Housing Start	0.49	0.03	0.00	0.55	0.53	0.02	0.00	0.56
Real Consumption	0.49	0.03	0.00	0.48	0.63	0.03	0.00	0.62
Real residential investment	0.41	0.03	0.00	0.44	0.56	0.02	0.00	0.64
Real nonresidential investment	0.48	0.02	0.00	0.49	0.61	0.03	0.00	0.61
Real state and local government consumption	0.43	0.04	0.00	0.40	0.60	0.05	0.00	0.56
Real federal government consumption	0.47	0.04	0.00	0.48	0.62	0.03	0.00	0.62
Unemployment rate	0.49	0.02	0.00	0.54	0.56	0.02	0.00	0.62
Three-month Treasury rate	0.55	0.02	0.00	0.59	0.63	0.03	0.00	0.67
Ten-year Treasury rate	0.51	0.02	0.00	0.54	0.60	0.02	0.00	0.63
AAA Corporate Rate Bond	0.54	0.02	0.00	0.56	0.61	0.02	0.00	0.62

Notes: The table shows the result from regression 2.9. Columns (1)-(3) report coefficients, standard errors and p-values from the panel data regression with time and individual fixed effect. Column (4) reports the median coefficient from individual regressions. Columns (5)-(8) reports the same statistics for the 2 quarters horizon. Standard errors are robust to heteroskedasticity and clustered at both forecaster and time level.

$G_1^2 > 0$ and the difference is positive: the gain implied by the CG regression overestimate the actual gain in presence of public information (or underestimate the stickiness).

Table 2.4 reports the estimated gain G from regression 2.9 in columns (1) -(2) and the gain G_{CG} implied by CG estimate 2.4 in absence of public information in columns (3)-(4). Figure 2.1 show the comparison graphically. Our estimate gain is less volatile across variables and consistently lower than the one implied by CG. We report the difference $G_{CG} - G$ in Table 2.4, columns (5)-(6), and plot it graphically in figure 2.2. The difference is consistently positive as implied by proposition 2. We test whether the difference is statistically larger than zero and report the p-value in column (7). The null hypothesis is rejected at the 10% confidence level for 7 variables out of 15.

The evidence indicate that public information is in fact an important part of the information set of forecasters. While the CG estimate implies very different gains

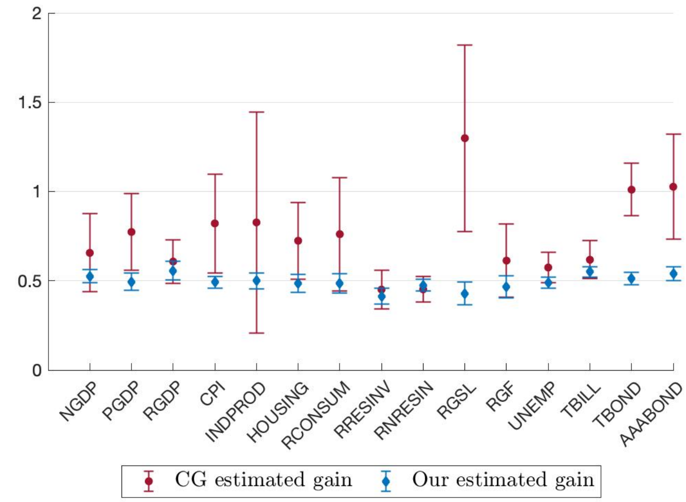


Figure 2.1: Estimated gains with the two methods

Notes: this figure plots the coefficient from panel data regression 2.13 with individual fixed effect. The blue diamonds represent the coefficient β_1 while the red circles represent the coefficient β_2. Standard errors are robust and clustered at the individual forecaster level.

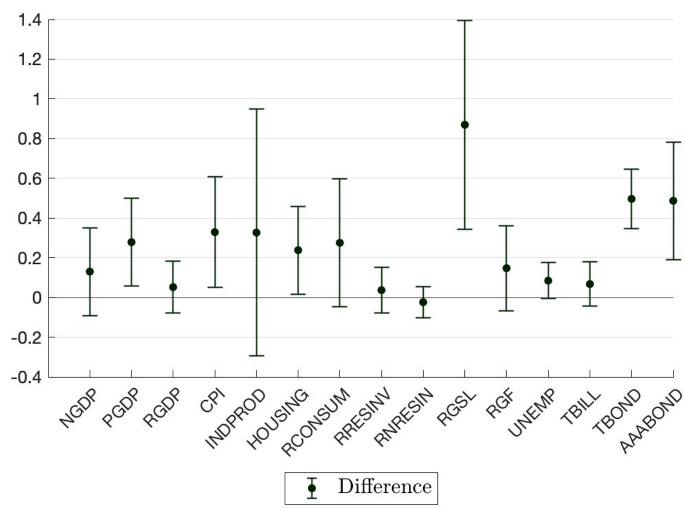


Figure 2.2: Difference between the two estimates

Notes: this figure plots the coefficient from panel data regression 2.13 with individual fixed effect. The blue diamonds represent the coefficient β_1 while the red circles represent the coefficient β_2. Standard errors are robust and clustered at the individual forecaster level.

Table 2.4: Difference between estimated gains

Variable	G_{CG} (1)	SE (2)	G (3)	SE (4)	Difference (5)	SE (6)	p-value (7)
Nominal GDP	0.66	0.13	0.53	0.02	0.13	0.13	0.17
GDP price index inflation	0.77	0.13	0.49	0.03	0.28	0.13	0.02
Real GDP	0.61	0.07	0.56	0.03	0.05	0.08	0.26
Consumer Price Index	0.82	0.17	0.49	0.02	0.33	0.17	0.03
Industrial production	0.83	0.38	0.50	0.03	0.33	0.38	0.19
Housing Start	0.72	0.13	0.49	0.03	0.24	0.13	0.04
Real Consumption	0.76	0.19	0.49	0.03	0.28	0.20	0.08
Real residential investment	0.45	0.07	0.41	0.03	0.04	0.07	0.30
Real nonresidential investment	0.45	0.04	0.48	0.02	-0.02	0.05	0.69
Real state and local government consumption	1.30	0.32	0.43	0.04	0.87	0.32	0.00
Real federal government consumption	0.61	0.12	0.47	0.04	0.15	0.13	0.13
Unemployment rate	0.57	0.05	0.49	0.02	0.08	0.05	0.06
Three-month Treasury rate	0.62	0.07	0.55	0.02	0.07	0.07	0.16
Ten-year Treasury rate	1.01	0.09	0.51	0.02	0.50	0.09	0.00
AAA Corporate Rate Bond	1.03	0.18	0.54	0.02	0.49	0.18	0.00

Notes: Columns (1)-(2) reports the implied gain from CG regressions of table 2.2. Columns (3)-(4) replicate the gain estimate in table 2.3. Columns (5)-(8) reports the difference between column (1) and (3), its standard error and the probability of rejecting the null of column (5) lower or equal to zero.

across variables, with some series with no apparent stickiness (Ten-year Treasury rate, AAA Corporate Bond and Real federal government consumption), our novel approach suggests that the overall gain on new information is instead similar across variables but with differences in the role of public information. In particular, public exceed private information in importance for financial variable, consistently with the idea that most of private information is priced in an efficient market.

2.2.3 Fact 3: over-reaction in individual forecast

Individual forecast error is defined as the actual realization minus the individual forecast: $fe^i_{t+h,t} = x_{t+h} - \tilde{E}^i_t[x_{t+h}]$. Similarly, individual forecast revision is defined as the individual forecast provided today minus the forecast provided in the previous period about the same horizon: $fr^i_{t+h,t} = \tilde{E}^i_t[x_{t+h}] - \tilde{E}^i_{t-1}[x_{t+h}]$.

Bordalo et al. (2018b), hereafter BGMS, test the rational expectation hypothesis by regressing individual forecast errors on consensus forecast revisions. Intuitively, if forecasters were fully rational, it should not be possible to predict individual future errors using today individual revisions, which would be part of the forecasters' information sets. They run the following regression

$$fe^i_{t+h,t} = \alpha + \beta_{BMGS} fr^i_{t+h,t} + err^i_t \tag{2.11}$$

Proposition 6 *If agents forecasts follow 2.3, the coefficient of regression 2.11 is given by:*

$$\beta_{BGMS} = \frac{1-G}{G} - \frac{\frac{G_1^2}{K}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}{G^2\tilde{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}} \tag{2.12}$$

with $\tilde{\Sigma} == \equiv var(x_t - \tilde{E}^i_{t-1}[x_t]) = \frac{\rho^2[G_1^2\nu^{-1} + G_2^2\tau^{-1}] + \xi^{-1}}{1-\rho^2(1-G)^2}$.

Corollary 8 *If agents forecasts follow 2.3, under rational expectation the coefficient* β_{BGMS} *of regression 2.11 is equal to zero.*

According to RE, individual forecast errors should not be predictable using individual forecast revisions. A positive $\beta_{BMGS} > 0$ would imply that after an positive surprise today agents don't update their forecast enough and they consistently underestimate the future value of x. On the opposite, a negative $\beta_{BMGS} < 0$ would imply that after an positive surprise today agents become too optimistic and they consistently overestimate the future value of x.

BMGS documents a robust $\beta_{BMGS} < 0$ for a wide range of macroeconomic and financial series. We replicate their panel data econometric specification with individual fixed effects in columns 1-3 of Table 2.6. However, the panel specification could introduce a bias in the coefficient.[3] Therefore we present also the median coefficients from the individual level regressions in the last column of Table 2.6, which confirms the panel results.

[3] RE implies that it is not possible for agents to predict their own forecast errors, $\beta_{BGMS} = 0$. However since the panel regression exploits the cross sectional variance in addition to the time series one, this specification effectively uses the average information set to pin down β_{BMGS} and not only the individual one.

BGMS documents that overreaction holds also under the assumption that the fundamental process follows an AR(2). We replicate their finding in table B.2 in appendix B.4.

2.2.4 Fact 4: under-reaction to public information in individual forecast

Motivated by the importance of public information in forecasters information set, we differentiate individual forecast reaction to private and public information. We document that while individual forecasts overreact to new information in general, they underreact to new public information.

In order to measure public information, we use the lagged consensus forecast, namely the average of the individual forecasts provided in the previous quarter about the same horizon. The consensus forecast is available to the forecasters at the time of the survey. To capture the surprise component in the public information, we compute the difference between the public signal and individual prior about the signal: $pi_{t,t+h} = g_t - \tilde{E}^i_{t-1}[x_{t+h}]$. We run the following regression:

$$fe^i_{t+h,t} = \alpha + \beta_1 fr^i_{t+h,t} + \beta_2 pi_{t+h,T} + err^i_t \tag{2.13}$$

where g_t is a public signal providing information about the variable at horizon $t+h$.

Proposition 7 *If agents forecasts follow 2.3, the coefficients of regression 2.13 are given by:*

$$\begin{aligned}
\hat{\beta}_1 &= \frac{1-G_2}{G_2} - \frac{(\tilde{\Sigma}+\nu^{-1})G_2\frac{1}{\tau}}{(\tilde{\Sigma}+\nu^{-1})(G^2\tilde{\Sigma}^{-1}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\tilde{\Sigma}+G_1\nu^{-1})^2} \\
\hat{\beta}_2 &= -\frac{G_1}{G_2} + \frac{(G\tilde{\Sigma}+G_1\nu^{-1})G_2\frac{1}{\tau}}{(\tilde{\Sigma}+\nu^{-1})(G^2\tilde{\Sigma}+G_1^2\nu^{-1}+G_2^2\tau^{-1})-(G\tilde{\Sigma}+G_1\nu^{-1})^2}
\end{aligned} \tag{2.14}$$

with $\tilde{\Sigma} \equiv var(x_t - \tilde{E}^i_{t-1}[x_t]) = \frac{\rho^2[G_1^2\nu^{-1}+G_2^2\tau^{-1}]+\xi^{-1}}{1-\rho^2(1-G)^2}$.

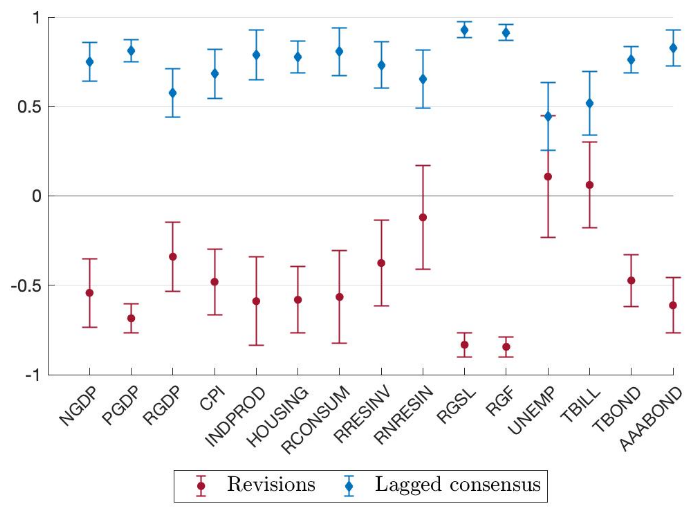


Figure 2.3: Forecast errors on forecast revisions and public information

Notes: this figure plots the coefficient from panel data regression 2.13 with individual fixed effect. The blue diamonds represent the coefficient β_1 while the red circles represent the coefficient β_2. Standard errors are robust and clustered at both time and individual forecaster level.

Corollary 9 *If agents forecasts follow 2.3, under rational expectation the coefficient β_1 and β_2 of regression 2.13 are equal to zero.*

Under RE and truthful revelation, it would not be possible to predict individual errors using individual information sets. Panel A of Table 2.5 reports the panel data regressions with individual fixed effects and the median from individual regressions. Both specifications display a consistent $\beta_1 < 0$ and $\beta_2 > 0$ across variables, with few exceptions, meaning individual overreaction to private information and underreaction to public information. Figure 2.3 provide a graphical representation of the estimated coefficient from the panel data regression. In table B.3 in appendix B.4 we show that these results holds also under the assumption that the fundamental follows an AR(2) process.

Discussion We generalized two known fact in the information literature, fact 1 and fact 3, with the inclusion of public information and document two new facts, fact 2 and 4, which provides new insight on how to model agent's belief formation.

While our third fact, overreaction to new information, seems to indicates a departure for the rational expectation hypothesis, we distinguish between new private and public information and find that forecasters overreact to the first but underreact to the second, our fourth fact. This is not consistent with model of overreaction to all new information as Bordalo et al. (2020) and Broer and Kohlhas (2018), but it is consistent with other two distinct frameworks. First, behavioral overconfidence, according to which agents overestimate the actual precision of their private signal. Second, strategic incentives, according to which agents are in fact rational in their beliefs formation, but provide to the public biased forecast in which their overweight private signals in order to stand out from the crowd.

While the behavioral overconfidence model departs from rational expectation, the strategic incentives model departs from truthful revelation. Both of them are consistent with facts 1-4, but with an important difference. Overconfident agents believe the forecast they post, and fact 2 correctly identify the stickiness of new information update. However, strategic agents beliefs are different from what they report to the survey, and fact 2 underestimate the actual stickiness (overestimate the gain) of their honest beliefs. Since posted forecasts overweight new private information, they appear to be less sticky than actual beliefs, with two consequences: first, the consensus forecast is more accurate, as it averages out private noise; second, the forecast dispersion is larger than the honest dispersion.

In the remaining part of the paper, we provide a general theoretical framework of strategic incentives consistent with our four empirical facts and estimate it structurally, in order to recover the actual stickiness of belief updating and forecast dispersion.

2.3 Strategic incentives in Forecast Reporting

2.3.1 Static setting

In this section we present a model of strategic interactions in forecast reporting, in which forecasters don't only want to provide accurate forecasts, but also to stand out with respect to the average forecast. We therefore depart from the assumption of truthful reporting by introducing strategic substitutability between forecasters.

The likelihood of strategic interactions in reporting is known in the forecasting literature. For example, Croushore (1997) suggests that "some [survey] participants might shade their forecasts more toward the consensus (to avoid unfavorable publicity when wrong), while others might make unusually bold forecasts, hoping to stand out from the crowd".[4] Ottaviani and Sørensen (2006) model the latter interpretation in a winner-take-all game. While we also focus on strategic substitutability, we keep a more general Morris and Shin (2002) global game setting.

In particular, the forecasters' problem is

$$min_{\hat{x}^i} \quad u^i = E^i\left[(\hat{x}^i - x)^2 - \lambda(\hat{x}^i - \bar{\hat{x}})^2\right] \tag{2.15}$$

where x is the true state and $\bar{\hat{x}} = \int \hat{x}^i di$ is the average of the reported forecast $\hat{x}^i$; $0 < \lambda < 1$ measures the degree of strategic substitutability in agent's reported forecasts.

The first order condition is:

$$\hat{x}^i = \frac{1}{1-\lambda}E^i[x] - \frac{\lambda}{1-\lambda}E^i[\bar{\hat{x}}] \tag{2.16}$$

If $\lambda = 0$, agents report their honest beliefs. If $\lambda > 0$, agents not only want to be

[4] While strategic considerations apply more intuitively to non-anonymous survey, in Appendix B.5 we argue that they apply to anonymous survey as well, as forecasters are likely to provide the same forecast to both surveys.

accurate, but also to stand out with respect to the average forecast.

Information Suppose the actual x is unobserved. Forecasters have a common prior $x \sim N(\mu, \chi^{-1})$. Moreover, they received a private and a public signal, both unbiased and centered around the true x with some noise.

$$\begin{aligned} s^i &= x + \eta^i \\ g &= x + e \end{aligned} \tag{2.17}$$

with $\eta^i \sim N(0, \tau^{-1})$ and $e \sim N(0, \nu^{-1})$.

The resulting honest posterior is

$$E^i[x] = \mu + \gamma_1(g - \mu) + \gamma_2(s^i - \mu) \tag{2.18}$$

with $\gamma_1 = \frac{\nu}{\tau+\nu+\chi}$, $\gamma_2 = \frac{\tau}{\tau+\nu+\chi}$.

Introduce now strategic substitutability in expectation reporting as in equation 2.16. We guess a linear solution, solve for the fixed point problem and we get

$$\hat{x}^i = \mu + \delta_1(g - \mu) + \delta_2(s^i - \mu) \tag{2.19}$$

where $\delta_2 = \frac{\gamma_2}{(1-\lambda)+\lambda\gamma_2}$, $\delta_1 = \frac{(1-\lambda)\gamma_1}{(1-\lambda)+\lambda\gamma_2}$, $1 - \delta_1 - \delta_2 = \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)+\lambda\gamma_2}$. In order to stand out from the crowd, the forecasters overweight new private information in his posted forecast with respect to his actual beliefs ($\delta_2 > \gamma_2$) and underweight new public information ($\delta_1 < \gamma_1$). At the same time, since the prior is common and new information partly private, the agent overweight new information as a whole ($1 - \delta_1 - \delta_2 < 1 - \gamma_1 - \gamma_2$).

Proposition 8 *In a strategic substitutability game as in 2.16, with $0 < \lambda < 1$, the* *coefficient of the individual regression 2.11 is given by:*

$$\beta_{BGMS} = \frac{-\lambda\tau\chi}{([(1-\lambda)\nu + \tau]^2 + (1-\lambda)^2\nu\chi)} \tag{2.20}$$

Thus $\beta_{BGMS} < 0$ if $\lambda > 0$.

If $\lambda = 0$, there is no strategic interaction between forecasters and they simply report their honest beliefs. In that case, $\beta_{BGMS} = 0$, as forecast errors are not correlated with any information available in time t, and in particular her forecast revisions. This result follows directly from rational expectation. On the other hand, if $\lambda > 0$, agents overweight private information to stand out from the crowd, which results in $\beta_{BGMS} < 0$, meaning overreaction to new information. Ottaviani and Sørensen (2006) derive a similar result in a specific winner-take-all game only considering private information. The model reconciles the empirical result in section 2.2.

Proposition 9 *In a strategic substitutability game as in 2.16, with $0 < \lambda < 1$, the coefficient of the consensus regression 2.4 is given by:*

$$\beta_{CG} = \frac{(1-\lambda)\tau\chi}{([(1-\lambda)\nu + \tau]^2 + [(1-\lambda)^2\nu + \tau]\chi)} \tag{2.21}$$

Thus $\beta_{CG} > 0$ if $\lambda < 1$.

If $\lambda = 0$, there is no strategic interaction between forecasters and they simply report their honest beliefs. In that case, $\beta_{CG} > 0$: the average forecast is sub-optimally sticky with respect to the average signal, which is less noisy than the individual one as shown by CG. On the other hand, if $\lambda > 0$, agents overweight new information and the average forecast is less sticky. The higher is the strategic incentive λ, the lower is the rigidity of posted forecast. In the limit case of $\lambda \to 1$, individual forecasters adjust one-to-one their posteriors to new information, making the average forecast not sticky. It is not possible to have $\beta_{CG} < 0$,consistently with the data.

Proposition 10 *In a strategic substitutability game as in 2.16, with $0 < \lambda < 1$, the coefficient of the individual regression 2.13 is given by:*

$$\begin{aligned}\beta_1 &= \frac{-\lambda(\nu+\chi)}{(\tau+\nu+\chi)} \\ \beta_2 &= \frac{\lambda\nu}{(\tau+\nu+\chi)}\end{aligned} \tag{2.22}$$

Thus $\beta_1 > 0$ and $\beta_2 < 0$ if $\lambda > 0$.

Proposition 7 represents our main theoretical result. If $\lambda = 0$, forecasters report their honest beliefs and both $\beta_1 = 0$ and $\beta_2 = 0$ as implied by rational expectation. However, with strategic incentives $\lambda > 0$, forecasters overweight private information and underweight public information, in order to stand from the crowd. This leads to an underreaction to public information, as measured by $\beta_2 > 0$, and overreaction to private information, as measured by $\beta_1 < 0$. This result reconciles our new empirical fact four documented in section 2.2.

2.3.2 Dynamic setting

We now extend the previous strategic incentives model to a dynamic setting. Assume the series follows a AR(1) process

$$x_t = \rho x_{t-1} + u_t \tag{2.23}$$

with $u \sim N(0, \xi^{-1})$.

Each agent receive a private signal s_t^i and a public signal g_t

$$\begin{aligned}s_t^i &= x_t + \eta_t^i \\ g_t &= x_t + e_t\end{aligned} \tag{2.24}$$

with $\eta_t^i \sim N(0, \tau^{-1})$, $e_t \sim N(0, \nu^{-1})$.

Honest beliefs Agents form beliefs about x at horizon h: $E_t^i[x_{t+h}] \equiv x_{t+h,t}^i$. The honest posterior belief about x is given by the Kalman filter

$$x_{t,t}^i = x_{t,t-1}^i + K_{1,1}(g_t - x_{t,t-1}^i) + K_{1,2}(s_t^i - x_{t,t-1}^i)$$

where the Kalman gains are

$$\begin{aligned} K_{1,1} &= \frac{\nu}{\Sigma^{-1} + \nu + \tau} \\ K_{1,2} &= \frac{\tau}{\Sigma^{-1} + \nu + \tau} \end{aligned} \tag{2.25}$$

and the posterior forecast error variance

$$\begin{aligned} \Sigma &\equiv E[(x_t - x_{t,t-1}^i)(x_t - x_{t,t-1}^i)'] \\ &= \frac{-[(\rho^2 - 1)\xi + (\tau + \nu)] + \sqrt{[(\rho^2 - 1)\xi + (\tau + \nu)]^2 + 4(\tau + \nu)\xi}}{2} \end{aligned} \tag{2.26}$$

Strategic interactions As in the previous section, the strategic substitutability in agents objective function leads them to report

$$\begin{aligned} \hat{x}_{t,t}^i &= \frac{1}{1-\lambda} x_{t,t}^i - \frac{\lambda}{1-\lambda} E^i[\bar{\hat{x}}_{t,t}] \\ \hat{x}_{t+h,t}^i &= \rho^h \hat{x}_{t,t}^i \end{aligned} \tag{2.27}$$

where $\hat{x}_{t+h,t}^i$ is the forecast provided by individual i in t about realization in $t+h$, and $\bar{\hat{x}}_{t+h,t} = \int^i \hat{x}_{t+h,t}^i di$ is the average of forecasts provided in t about realization in $t+h$. If $\lambda = 0$, agents report their true beliefs. With $0 < \lambda < 1$, agents not only want to be accurate, but also to stand out with respect to the average forecast.

The model builds on Woodford (2001) and Coibion and Gorodnichenko (2012).[5]

[5] Our model depart from the latter in two dimensions. First, while they consider only strategic complementarity, we focus on strategic substitutability. Second, while they only consider consensus forecasts, we are interested in individual forecasts

Following them, we average $\hat{x}^i_{t,t}$ across agents and use repeated substitution in 2.27 to express the reported average forecast as

$$F_t = -\frac{1}{1-\lambda}\sum_{k=0}^{\infty}\left(\frac{\lambda}{1-\lambda}\right)^k \bar{E}^{(k)}[x_t] = \frac{1}{1-\lambda}\bar{x}_{t+h,t} - \frac{\lambda}{1-\lambda}\bar{E}_t[\bar{\hat{x}}_{t+h,t}] \tag{2.28}$$

We guess and verify the law of motion for F_t and the other unobserved state variables. In particular, we conjecture that the state vector evolves according to[6]

$$Z \equiv \begin{bmatrix} x_t \\ F_t \\ w_t \end{bmatrix} = MZ_{t-1} + m \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{2.29}$$

Where

$$M = \begin{bmatrix} \rho & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad m = \begin{bmatrix} 1 & 0 \\ m_{2,1} & m_{2,2} \\ 0 & 1 \end{bmatrix} \tag{2.30}$$

the observable variables are the two signals about x_t

$$V^i_t \equiv \begin{bmatrix} g_t \\ s^i_t \end{bmatrix} = HZ_t + \begin{bmatrix} 0 \\ \eta^i_t \end{bmatrix} \tag{2.31}$$

where

$$H = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \tag{2.32}$$

Agents use their conjecture law of motion 2.29 and the observables 2.31 to infer the state using the individual Kalman filter. The posterior estimate of the state

[6] w_t takes care of the correlation between public signal and higher order beliefs F_t

vector by agent i is

$$\begin{aligned} E_t^i[Z_t] &= ME_{t-1}^i[Z_{t-1}] + K(V_t^i - E_{t-1}^i[V_t]) \\ &= (I-KH)ME_{t-1}^i[Z_{t-1}] + KHMZ_{t-1} + KHm\begin{bmatrix} u_t \\ e_t \end{bmatrix} + K\begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \end{aligned} \tag{2.33}$$

Where K is the Kalman gain. Average 2.33 to find the consensus believe on the state vector.

$$\bar{E}_t[Z_t] = (I-KH)M\bar{E}_{t-1}[Z_{t-1}] + KHMZ_{t-1} + KHm\begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{2.34}$$

From the definition on F_t in 2.28 it follows that

$$\begin{aligned} F_t &= \begin{bmatrix} \frac{1}{1-\lambda} & -\frac{\lambda}{1-\lambda} & 0 \end{bmatrix} \bar{E}_t[Z_t] \equiv \xi\bar{E}_t[Z_t] \\ &= \xi(I-KH)M\bar{E}_{t-1}[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm\begin{bmatrix} u_t \\ e_t \end{bmatrix} \end{aligned} \tag{2.35}$$

$$\begin{aligned} F_t =&((1-\alpha)\rho + \alpha G)\bar{E}_{t-1}[x_{t-1}] + \alpha L\bar{E}_{t-1}[F_{t-1}] - C\rho\bar{E}_{t-1}[x_{t-1}] \\ &+ C\rho x_{t-1} + C_1e_t + Cu_t \\ =&[\rho(1-\alpha) + \alpha G - (1-\alpha)L - C\rho]\bar{E}_{t-1}[x_{t-1}] + \\ &+ LF_{t-1} + C\rho x_{t-1} + Cu_t + C_1e_t \end{aligned} \tag{2.36}$$

where we used 2.28 to substitute

$$\alpha\bar{E}_t[F_{t-1}] = F_{t-1} - (1-\alpha)\bar{E}_{t-1}[x_{t-1}]$$

and we defined

$$C_1 \equiv \frac{K_{1,1} - \lambda(K_{2,1})}{1-\lambda}, \qquad C_2 \equiv \frac{K_{1,2} - \lambda K_{2,2}}{1-\lambda} \qquad \text{and} \qquad C = C_1 + C_2$$

Equation 2.36 must equal the second line of the perceived law of motion 2.29. The solution to the fixed point is given by $G = C\rho$, $m_{2,1} = C$, $m_{2,2} = C_1$ and $L = \rho - G$.

Given the law of motion of unobserved state 2.29 and the observable 2.31, the posterior variance of the forecast solves the following Ricatti equation

$$\begin{aligned} &\Sigma \equiv E[(Z_t - Z^i_{t,t-1})(Z_t - Z^i_{t,t-1})'] \\ &\Sigma = M(\Sigma - \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} H\Sigma)M' + m \begin{bmatrix} \xi^{-1} & 0 \\ 0 & \nu^{-1} \end{bmatrix} m' \end{aligned} \tag{2.37}$$

and the Kalman filter is

$$K = \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix} \right)^{-1} \tag{2.38}$$

Finally, the individual posted forecast is

$$\begin{aligned} \hat{x}^i_{t,t} &= \xi E^i_t[Z_t] \\ &= \hat{x}^i_{t,t-1} + C_1(g_t - \hat{x}^i_{t,t-1}) + C_2(s^i_t - \hat{x}^i_{t,t-1}) \end{aligned} \tag{2.39}$$

Note that the individual posted forecast updating in 2.39 is similar to individual Kalman Filter in 2.33, with C_1 and C_2 as ”modified” gains in place of K_1 and K_2. In particular, if $\lambda = 0$, $C_1 = K_1$ and $C_2 = K_2$: with no strategic incentives, agents simply report their honest beliefs. However, when $\lambda > 0$, one can show that $C_1 < K_1$ and $C_2 > K_2$: agents underweight public information and overweight private information in their posted forecast.[7]

The posted forecast updating 2.39 mirrors the general framework 2.3 in section 2.2 with $G_1 = C_1$, $G_2 = C_2$ and $\tilde{E}_t[x_t] = \hat{x}_{t,t}$. Therefore the coefficient from regressions 2.4, 2.9, 2.11 and 2.13 follows from propositions 1-4.

2.4 Structural estimation

We now proceed to estimate our model to test its ability to match the data recover the honest beliefs of forecasters to compute the actual belief stickiness. First, for each series we estimate the autoregressive coefficient ρ and the fundamental disturbance variance $\sigma_u^2 \equiv \xi^{-1}$ directly from the data. Then we use the simulated method of moments to estimate the remaining parameters of the model: the public noise variance $\sigma_e^2 \equiv \nu^{-1}$, the private noise variance $\sigma_\eta^2 \equiv \tau^{-1}$ and the strategic incentive parameter λ. We prefer the simulated method of moments to maximum likelihood as the simplicity of our model come at the cost of likely misspecification which would be problematic when using maximum likelihood.

The data moments we target are the mean cross sectional dispersion of forecast errors, the coefficient β_1 from the public information regression 2.13 and the posted gain C from regression 2.9. We choose these three moments as they are differently affected by the three parameters to be estimated and therefore provide good identification.[8]

Table 2.7 reports the estimated parameters for each series, while table 2.8 reports targeted and untargeted moments in the model and in the data. While the

[7] To see this, note that $K_{1,1} < K_{2,1}$: intuitively, the public signal is more informative about the average forecast than about the actual state, because of the average belief depends also on the public noise. On the other hand, $K_{1,2} > K_{2,2}$: intuitively, the private signal is less informative about the average forecast than about the actual state, as the average forecast also depends on the public noise.

[8] While larger strategic incentive parameter λ increases posted gain C, an increase in either private or public noise decreases it. On the other hand, the coefficient β_1 decreases in private noise and strategic incentives but increase in public noise (see proposition 7). Finally, strategic incentives and public noise always increase mean dispersion, while private noise initially increases it and then decreases it.

model is able to match the targeted model in columns 1-6, it also does a good job in matching the untargeted moments in columns 7-12 for most of the series. The only exceptions are the CG coefficient for the financial series in column 7-8 and a general underestimation of the public information coefficient in column 11-12.

We use the model to recover the honest beliefs of forecasters behind the biased posted forecast, and compute the related moments. We compare them with the moments calculated on the posted forecast in table 2.9. First, columns 1-3 reports the weight on new information in posted forecast, in honest beliefs and their ratio. Hones beliefs are stickier than posted forecasts, as the latter overweight new information. The magnitude of the difference can be appreciated by looking at the difference between posted and honest consensus mean forecast error. Because of the individual overweight of new information, the average posted forecast is less sticky and therefore more accurate than the honest one. For some variable (nominal GDP, CPI, Housing Start, Ten-year and AAA bond rate) the honest error is more than 1.5 times the posted one. Finally, columns 7-8 compare honest and posted cross sectional dispersion of forecast errors, which is used as a proxy for uncertainty in the literature (e.g. Kozeniauskas et al. 2018). The overweight of private information increase substantially the dispersion of forecasts, as the honest beliefs dispersion is less than half the posted one for most of the series.

Table 2.5: Individual errors on revisions

	3 quarters horizon				2 quarters horizon			
	β_{BGMS}	SE	p-value	Median	β_{BGMS}	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.25	0.08	0.00	-0.19	-0.11	0.06	0.10	-0.08
GDP price index inflation	-0.35	0.04	0.00	-0.35	-0.25	0.04	0.00	-0.26
Real GDP	-0.10	0.08	0.22	0.07	-0.07	0.10	0.47	0.04
Consumer Price Index	-0.30	0.09	0.00	-0.29	-0.24	0.07	0.00	-0.24
Industrial production	-0.30	0.14	0.04	-0.31	-0.01	0.10	0.94	0.03
Housing Start	-0.28	0.09	0.00	-0.28	0.12	0.05	0.03	0.07
Real Consumption	-0.26	0.12	0.04	-0.24	-0.16	0.08	0.07	-0.16
Real residential investment	-0.08	0.10	0.44	-0.07	0.07	0.08	0.41	0.02
Real nonresidential investment	0.08	0.13	0.56	0.15	0.10	0.07	0.18	0.10
Real state and local government consumption	-0.56	0.05	0.00	-0.52	-0.30	0.05	0.00	-0.26
Real federal government consumption	-0.48	0.04	0.00	-0.40	-0.28	0.04	0.00	-0.27
Unemployment rate	0.24	0.16	0.13	0.19	0.20	0.11	0.09	0.20
Three-month Treasury rate	0.24	0.10	0.03	0.29	0.14	0.08	0.09	0.21
Ten-year Treasury rate	-0.22	0.07	0.01	-0.24	-0.24	0.09	0.01	-0.27
AAA Corporate Rate Bond	-0.27	0.07	0.00	-0.32	-0.22	0.06	0.00	-0.29

Notes: The table reports the coefficients from the BGMS regression (individual forecast errors on individual revisions). Columns (1) to (3) shows the panel data with fixed effect coefficient with standard errors and corresponding p-values. Standard errors are robust to heteroskedasticity and clustered at both forecaster and time level. Columns (7) shows the median coefficient of the BGMS regression at the individual level.

Table 2.6: Private and public information

Panel A: 3 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.54	0.12	0.00	-0.44	0.75	0.07	0.00	0.76
GDP price index inflation	-0.68	0.05	0.00	-0.64	0.81	0.04	0.00	0.83
Real GDP	-0.34	0.12	0.01	-0.18	0.58	0.08	0.00	0.62
Consumer Price Index	-0.48	0.11	0.00	-0.46	0.68	0.08	0.00	0.69
Industrial production	-0.59	0.15	0.00	-0.60	0.79	0.08	0.00	0.78
Housing Start	-0.58	0.11	0.00	-0.53	0.78	0.05	0.00	0.71
Real Consumption	-0.57	0.16	0.00	-0.58	0.81	0.08	0.00	0.81
Real residential investment	-0.38	0.15	0.01	-0.39	0.73	0.08	0.00	0.66
Real nonresidential investment	-0.12	0.18	0.50	-0.10	0.65	0.10	0.00	0.51
Real state and local government consumption	-0.83	0.04	0.00	-0.81	0.93	0.03	0.00	0.89
Real federal government consumption	-0.84	0.03	0.00	-0.77	0.91	0.03	0.00	0.87
Unemployment rate	0.11	0.21	0.61	-0.02	0.44	0.11	0.00	0.42
Three-month Treasury rate	0.06	0.15	0.68	0.11	0.52	0.11	0.00	0.38
Ten-year Treasury rate	-0.47	0.09	0.00	-0.40	0.76	0.04	0.00	0.83
AAA Corporate Rate Bond	-0.61	0.09	0.00	-0.67	0.83	0.06	0.00	0.87

Panel B: 2 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.35	0.09	0.00	-0.27	0.62	0.06	0.00	0.63
GDP price index inflation	-0.55	0.06	0.00	-0.50	0.70	0.04	0.00	0.66
Real GDP	-0.25	0.13	0.06	-0.14	0.54	0.08	0.00	0.54
Consumer Price Index	-0.38	0.09	0.00	-0.36	0.52	0.08	0.00	0.52
Industrial production	-0.16	0.12	0.19	-0.14	0.49	0.08	0.00	0.50
Housing Start	-0.15	0.08	0.08	-0.15	0.54	0.05	0.00	0.56
Real Consumption	-0.37	0.11	0.00	-0.29	0.64	0.07	0.00	0.69
Real residential investment	-0.13	0.11	0.23	-0.16	0.49	0.07	0.00	0.43
Real nonresidential investment	-0.02	0.10	0.81	-0.04	0.41	0.07	0.00	0.44
Real state and local government consumption	-0.63	0.08	0.00	-0.51	0.79	0.04	0.00	0.72
Real federal government consumption	-0.71	0.06	0.00	-0.64	0.80	0.04	0.00	0.74
Unemployment rate	0.09	0.15	0.57	0.03	0.39	0.10	0.00	0.37
Three-month Treasury rate	0.02	0.11	0.89	0.10	0.48	0.10	0.00	0.39
Ten-year Treasury rate	-0.46	0.11	0.00	-0.45	0.71	0.07	0.00	0.67
AAA Corporate Rate Bond	-0.49	0.08	0.00	-0.52	0.70	0.06	0.00	0.71

Notes: this table reports the coefficients of regression 2.13 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

Table 2.7: Estimasted parameters

Variable	ρ (1)	$\frac{\sigma_e}{\sigma_u}$ (2)	$\frac{\sigma_\eta}{\sigma_u}$ (3)	λ (4)
Nominal GDP	0.93	1.48	1.70	0.74
GDP price index inflation	0.93	1.60	2.13	0.88
Real GDP	0.80	1.30	1.36	0.47
Consumer Price Index	0.78	1.38	1.60	0.61
Industrial production	0.85	1.28	1.86	0.68
Housing Start	0.85	1.38	1.81	0.70
Real Consumption	0.87	1.33	1.84	0.67
Real residential investment	0.89	1.56	1.74	0.49
Real nonresidential investment	0.89	2.37	1.28	0.25
Real state and local government consumption	0.89	1.32	2.79	0.90
Real federal government consumption	0.80	1.29	2.90	0.87
Ten-year Treasury rate	0.83	1.81	1.56	0.72
AAA Corporate Rate Bond	0.85	1.76	1.82	0.87

Table 2.8: Moments in data and model

	Targeted moments						Untargeted moments					
	Mean Dispersion		C		β_1		β_{CG}		β_{BGMS}		β_2	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	1.49	1.49	0.53	0.53	-0.54	-0.54	0.52	0.41	-0.25	-0.31	0.75	0.21
GDP price index inflation	0.33	0.33	0.49	0.49	-0.68	-0.68	0.29	0.50	-0.35	-0.44	0.81	0.31
Real GDP	0.92	0.92	0.56	0.56	-0.34	-0.34	0.65	0.33	-0.10	-0.15	0.57	0.13
Consumer Price Index	0.31	0.31	0.49	0.49	-0.48	-0.48	0.22	0.38	-0.30	-0.24	0.67	0.16
Industrial production	3.71	3.71	0.50	0.50	-0.59	-0.59	0.21	0.22	-0.30	-0.22	0.79	0.26
Housing Start	110.04	110.04	0.49	0.49	-0.58	-0.58	0.38	0.32	-0.28	-0.28	0.78	0.23
Real Consumption	0.51	0.51	0.49	0.49	-0.56	-0.56	0.31	0.25	-0.26	-0.23	0.80	0.23
Real residential investment	27.03	27.03	0.41	0.41	-0.37	-0.37	1.22	0.40	-0.08	-0.17	0.73	0.11
Real nonresidential investment	7.38	7.38	0.48	0.48	-0.12	-0.12	1.21	0.94	0.08	-0.10	0.65	0.01
Real state and local government consumption	1.41	1.41	0.47	0.47	-0.84	-0.84	0.63	0.17	-0.48	-0.41	0.91	0.45
Real federal government consumption	6.40	6.40	0.43	0.43	-0.83	-0.83	-0.23	0.12	-0.56	-0.35	0.93	0.37
Ten-year Treasury rate	0.17	0.17	0.51	0.51	-0.47	-0.47	-0.01	0.69	-0.22	-0.38	0.76	0.09
AAA Corporate Rate Bond	0.34	0.34	0.54	0.54	-0.61	-0.61	-0.03	0.62	-0.27	-0.48	0.83	0.18

Table 2.9: Posted and honest moments

	Gain			Consensus MSE			Dispersion		
	Posted	Honest	Ratio	Posted	Honest	Ratio	Posted	Honest	Ratio
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nominal GDP	0.53	0.40	0.76	0.49	1.07	2.19	1.49	0.29	0.19
GDP price index inflation	0.49	0.32	0.66	0.05	0.14	2.92	0.33	0.02	0.06
Real GDP	0.56	0.49	0.88	0.78	1.14	1.47	0.92	0.41	0.44
Consumer Price Index	0.49	0.40	0.82	0.23	0.36	1.58	0.31	0.08	0.27
Industrial production	0.50	0.44	0.87	3.51	5.11	1.46	3.71	0.60	0.16
Housing Start	0.49	0.40	0.82	69.95	115.75	1.65	110.04	18.10	0.16
Real Consumption	0.49	0.42	0.86	0.46	0.68	1.49	0.51	0.09	0.18
Real residential investment	0.41	0.36	0.87	29.60	40.95	1.38	27.03	10.76	0.40
Real nonresidential investment	0.48	0.43	0.90	4.12	5.30	1.29	7.38	6.01	0.82
Real state and local government consumption	0.47	0.40	0.86	0.54	0.81	1.51	1.41	0.02	0.02
Real federal government consumption	0.43	0.39	0.90	5.96	7.49	1.26	6.40	0.14	0.02
Ten-year Treasury rate	0.51	0.33	0.64	0.04	0.11	2.55	0.17	0.05	0.27
AAA Corporate Rate Bond	0.54	0.29	0.54	0.04	0.14	3.75	0.34	0.04	0.11

Chapter 3

International Trade and Portfolio Diversification: the Role of Information

3.1 Introduction

A long standing puzzle in the international finance literature concerns the empirical positive relation between bilateral trade in goods and portfolio investment between countries (Portes and Rey, 2005; Aviat and Coeurdacier, 2007; Lane and Milesi-Ferretti, 2008; Galstyan et al., 2016). While this finding is hard to reconcile with risk hedging motives, I show that it can be rationalized in an investment model with endogenous information.

According to a standard risk-hedging view, agents should diversify the positive correlation between financial and non-financial income by holding a higher share of foreign asset (Baxter and Jermann, 1997). Since trade makes domestic non-financial income more dependent on trading partner's risk factors, it follows that investors should diversify their income by holding fewer trade partner's equities as bilateral trade increases. This basic result is in stark contrast with the empirical positive relation between bilateral investment and trade across country pairs.

I show that allowing investors to collect information before their investment decision helps rationalize the empirical evidence. I follow the literature in introducing investors with preferences for early resolution of uncertainty (Van Nieuwerburgh and Veldkamp, 2009). I show that investors decide to collect information on trading partner's country risk factors to decrease the perceived riskiness of their non-financial income. As a result, they perceive the foreign country as less risky, leading to higher desired bilateral investment. I show that information choice decreases the magnitude of risk-hedging motives and can offset diversification incentives, leading to higher investment in trading country's assets. To the best of my knowledge, this is the first paper to investigate theoretically the relation the impact of information choice on investment and trade.[1]

Model I consider a two-country static investment model, where investors receive (i) financial income from their asset portfolio and (ii) non-financial income from wages. They decide the composition of their portfolio between a domestic asset, a foreign asset and a risk-free asset. They receive a wage from a local firm selling domestic goods in both countries. Preferences for domestic and foreign goods in each country determines the amount of trade and therefore firm's (and wage's) exposure to domestic and foreign demand. I assume that both demand for consumption goods and asset return in each country depends positively on the same country-specific risk factor, and therefore are positively correlated.

I show that this setting with exogenous information can not match the empirical evidence. In this model, the higher is trade with the foreign country, the more correlated is domestic investor's non financial income with foreign asset return, and therefore higher the incentive to not invest abroad in order to diversify income. Bilateral trade and portfolio investment are negatively correlated, contrary to the

[1] While also Dasgupta and Mondria (2018) applies the endogenous information approach to a trade model, they do not study investment, but jut trade flows.

data.

I show that a model with information choice leads to an opposite outcome with respect to the standard model and it is consistent with the evidence. Before investing, investors can acquire information on domestic and foreign countries with the intent to decrease perceived uncertainty of future financial and non-financial income.[2] While trade is given by preferences, investment decision depends on asset's perceived riskiness and therefore on information choice.

In absence of trade, domestic agent's non-financial income is affected only by domestic risk. Thus he collects information only on domestic risk factor and invest mostly on domestic country as he perceives it as less risky. When the amount of trade is larger, domestic agent's non-financial income is more exposed to foreign risk. Thus he collects information also on the foreign risk factor and increases investment in foreign asset as he perceives it as less risk. While the risk-hedging motives of the standard model are still present, I show that information choice decreases their importance in agent's portfolio allocation. Bilateral trade and portfolio investment are positively correlated, consistently with the data.

My model relies on the assumption that financial and non-financial returns are correlated, but it does not rely on a particular correlation sign.[3] If the correlation is positive, the mechanism works against risk hedging; if it is negative, it provides an amplification mechanism. Importantly, while a risk-hedging explanation always implies optimal risk sharing, the friction introduced by cognitive limitation makes the final allocation inefficient.

Empirical analysis The main implication of my model is that trade increases investment in assets through attention allocation. I test this implication in the data. I follow the financial literature in measuring attention using Google search queries

[2] I follow Van Nieuwerburgh and Veldkamp (2009) in using an Shannon (1948) entropy constraint to limit attention.

[3] One can think of it as a reduced form of country specific supply or demand shocks.

from *GoogleTrend*:[4] I proxy the attention of some country H to another country F with an index measuring the volume of Google researches in country H with the keyword ”F” in category ”finance”. I regress this proxy for attention on bilateral trade (export plus import), equity holding by households in origin country and other standard bilateral controls.[5] I have a panel of nearly 40 origin and 130 destination countries, from 2004 to 2015. Observations are at the country-pair level, and I include origin and destination country-time fixed effect.

My model implies that, controlling for equity holdings, trade affects positively attention allocation. The empirical test confirms my result: the impact of trade on the attention index is meaningful and strongly significant. I run some robustness tests using instrumental variables for trade and total assets to measure portfolio holdings, which confirm this result.

Contribution to the literature This paper contributes to different strands of the international macroeconomic and financial literature. First, it relates to the works on portfolio under-diversification and risk-hedging. Some papers explain observed equity home bias with real exchange rate risk hedging (Obstfeld and Rogoff, 2000; Coeurdacier, 2009; Benigno and Nistico, 2012)[6], while others argue that non-financial income and domestic equity returns are actually negatively correlated (Bottazzi et al., 1996; Julliard, 2002; Heathcote and Perri, 2013; Coeurdacier and Gourinchas, 2016). The most related to my paper is Heathcote and Perri (2013), which provides a model where domestic equity return are negatively correlated with non-financial income, but this covariance is less negative the more is the country's openness to trade. Therefore higher trade leads to higher diversification. This

[4] Da et al. (2011) and Andrei and Hasler (2014), while Mondria et al. (2010) use a different but similar proxy.

[5] I consider only households and not total holdings in order to address the objection that information frictions might not result from endogenous attention choices but they might be caused by private information that trading firms obtain with personal and business contacts in the foreign country.

[6] However, Van Wincoop and Warnock (2010) documents that the empirical correlation between exchange rate and equity return is too low to justify the observed home bias in equity.

paper provide a mechanism matching the same empirical finding without relying on particular correlation sign between financial and non-financial income, but with very different implication in terms of risk sharing efficiency.

Second, this paper relates to the empirical literature on international capital flows, documenting the positive relation between trade in goods and equity investment at the cross country level (Lane, 2000; Heathcote and Perri, 2013) and at the paired countries level (Portes and Rey, 2005; Aviat and Coeurdacier, 2007; Lane and Milesi-Ferretti, 2008; Galstyan et al., 2016). My contribution is to provide a theoretical model to rationalize this pattern. This paper is also consistent with Massa and Simonov (2006), which uses Swedish micro-data to argue that investment pattern does not seem to be explained by hedging risk, but by a "familiarity" effect.

Third, this paper relates to the works on portfolio home bias and information frictions (Merton, 1987; Gehrig, 1993; Brennan and Cao, 1997), and in particular later models of endogenous information (Van Nieuwerburgh and Veldkamp, 2009; Mondria et al., 2010; Valchev, 2017). I extend the investment model with endogenous information to incorporate non-financial income from trade and show that the structure of non-financial income affect final attention and portfolio allocation. Moreover, I show that the importance of the risk-hedging term depends on information choice.

The paper is organized as follows. Section 3.2 present some motivational evidence documenting the positive causal relation between trade in goods and portfolio investment between countries; section 3.3 develops an endogenous information model of investment with trade between two countries; section 3.4 presents the model solution; section 3.5 explains the results; section 3.6 brings the model's implications to the data; finally, section 3.7 concludes.

3.2 Motivating Evidence

This paper investigate the puzzling positive correlation between portfolio investment and trade between country pairs. The literature has widely documented this positive correlation using gravity-like equations (Portes and Rey, 2005; Aviat and Coeurdacier, 2007; Lane and Milesi-Ferretti, 2008; Galstyan et al., 2016). Moreover, the model presented here is consistent with two additional findings. First, Aviat and Coeurdacier (2007) document a strong positive causal relation from trade in goods to equity flows, but not the reverse direction. In line with this result, in my model trade will originate from consumers' preferences and affect portfolio investment.[7] Second, Galstyan et al. (2016) break down the results by holding sectors and find that gravity patterns in equity investment are weaker for professional investors than for retail ones. This finding is consistent with a model of cognitive limitations where professional investors are less cognitively bounded than retail investors.

I document empirically the positive relation between trade and portfolio investment as motivational fact for my theoretical contribution. First, I isolate the impact of trade on equity in the aggregate sample. Then, I consider only a subset of holders, the households.

My main regression is the following:

$$ln(Equity_{ijt}) = \phi_{it} + \phi_{jt} + \beta_1 ln(Trade_{ijt}) + \beta_3 X_{ijt} + \epsilon_{ijt} \tag{3.1}$$

The dependent variable is the total amount of country j equity in country i's portfolio *(IMF Coordinate Portfolio Investment Survey)*; the independent variable is the total amount fo trade between the two countries, measures as export plus import *(IMF Direction of Trade Statistics)*; all time varying country-specific characteristics

[7] Even though Aviat and Coeurdacier (2007) do not take a stand on why trade in goods causes equity investment, they mention the role of trade in decreasing information asymmetry as one possible explanation.

are controlled by source country-time fixed effect ϕ_{jt} and destination country-time fixed effect ϕ_{it}; In line with the literature, I add a number of bilteral controls: cultural (common language, colonial relationship), distance between capitals, common monetary union, free trade area, GDP correlation *(CEPII)*. The sample cover the period 2005-2015 for around 50 source and 150 destination countries. I exclude small financial centers and countries with less than one million inhabitants. The appendix reports a complete list of variables description and countries in the sample. The errors are clustered at the country-time level.

Column (1) of Table 3.1 reports the result using a simple OLS estimator. Trade has a positive and statistically significant impact on equity holding: a one percent increase in the bilateral amount of trade raises the bilateral equity holding of around 0.8 percent. This is the finding motivating this paper. The overall results is consistent with the previous literature.

In order to address a concern of possible reverse direction of causality, namely from asset holding to trade, I use a IV approach to isolate the exogenous effect of trade on portfolio investment. I again follow the literature (Aviat and Coeurdacier, 2007) in instrumenting trade with (i) dummy for free trade area, and (ii) bilateral trade costs. Column (2) reports the result. The test of weak identification, under-identification and the Hansen test ensures the instruments' validity. Trade's impact on equity holding is still positive, even if a bit lower in magnitude, confirming the previous results.

Column (3) and (4) repeats the same exercise but substituting equity holding with total holdings, meaning equity plus bond. The results are again confirmed. Table 3.2 reports an additional robustness check where I substitute trade with export from country i to country j, yielding the same results.

In the next section, I shed light on this finding by proposing a endogenous information model where trading agent optimally decide to be more informed about each other. However, a simpler explanation could be the following: if firms are trading with a foreign country, they might have access to information at a lower cost with respect to firms in non-trading countries, by physically being abroad or through business contacts. In order to address this concern, I run the same regression but now considering only the direct equity holding of households, which are less likely to have any information advantage.[8]

Table 3.3 mirrors the analysis of Table 3.1, but only considering household's holdings. The sample decreases considerably, but the impact of trade is pretty stable with respect to the previous tables. Table 3.4 reports the same results but substituting trade with export. The results are robust to this specification as well.

Motivated by this empirical fact, in the next section I develop a theoretical model able to formalize the link between trade and portfolio diversification through an information channel.

3.3 Model

In this section, I present a model in which trading agents jointly decide investment and amount of information. I show how the presence of trade affects investor's optimal choices: information and portfolio investment on foreign assets depends positively on the share of foreign good in agent's consumption.

My contribution extends the baseline two-countries model of investment and endogenous information (Van Nieuwerburgh and Veldkamp, 2009, 2010). Preferences for early resolution of uncertainty in this model lead to increasing returns to information and large amount of portfolio specialization with a even small initial information advantage. A simple intuition is that agents can expose most of their consumption to one of the two assets while decreasing its perceived volatility as

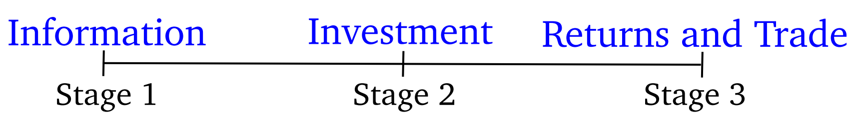
[8] The IMF Coordinated Portfolio investment Survey provide the data disaggregated by issuer and holder. Here I am keeping the aggregate level of issuer, but restricting the holder to household, therefore excluding firms, banks, government, mutual funds, et cetera.

much as possible, increasing by consequence its risk-adjusted return.

I extend this setting by including non-financial income and trade between countries. The new risk hedging component in the optimal portfolio would require to invest less in trading partner's asset to diversify risk. However, I show that investors want also to collect more information on trade partner's country to decrease non-financial income risk, and therefore want to invest more in its asset. Under certain parametrizations, the latter force is stronger than the former and the model implies higher portfolio investment in trading partner's asset.

3.3.1 Model structure

The model feature two countries, Home and Foreign, each one with a continuum of investors of measure ½. They face an investment choice between three assets: (i) a domestic risky asset h, (ii) a foreign risky asset f and (iii) a risk-free asset b. The model is static and divided in three stages:



1. **Information acquisition:** Agents face an attention allocation problem subject to a Shannon entropy contraint. [9] Each investor has a limited amount of attention they use to increase the precision of two signals they will receive in the second stage. These signals convey information on country-specific risk factors.

2. **Investment choice:** Each agent receives the signals and forms posteriors about domestic and foreign country risk factors. They choose how to allocate their initial resources to domestic, foreign and risk free assets.

[9] The information choice problem is in the spirit of the rational inattention literature (Sims, 2003)

3. **Portfolio returns and trade:** All shocks realize, agents receive the returns from portfolio investment and non-financial income. The latter is a wage compensation from working for a firm selling to domestic and foreign consumers.

I describe now each stage of the model backwards.

3.3.2 Third stage

In the third stage all shocks are realized and agents receive the returns from portfolio investment allocated in the second stage and their non-financial income. Final wealth is the sum of financial and non-financial wealth:

$$W_k = W_k^{NONFIN} + W_k^{FIN} \tag{3.2}$$

for $k \in \{h, f\}$. I describe each component of this income separately.

Non-financial income and trade

I model non-financial income as compensation from working for a firm selling domestic good domestically an abroad. Trade affect non-financial income by its impact on firm's aggregate demand's composition.

I assume that domestic and foreign demand for tradables are stochastic and depend on some country-specific risk factor. The implication is that investors receive income from tradable production, but do not consume tradables themselves. I make this assumption to maintain the model tractable and with the intent to take trade as given and study the impact on portfolio choices. Therefore in each countries there are three agents.

Investors In each country $k \in \{h, f\}$,, there is a continuum of investors with measure 1 that provides inelastically one unit of labor $L_k^{(i)}$ and receive a nominal wage $w_k^{(i)} L_k^{(i)} \equiv W^{NONFIN}$. Differently from the investment decision, labor supply is identical for each investor in the country, and they can be aggregated to a representative agent.

Firm In each country $k \in \{h, f\}$, a representative firm uses labor to produce a tradable good,with a production function $f(L)$ is linear in labor

$$Y_k = L_k \tag{3.3}$$

This firm face a domestic and a foreign demand for its good.

Consumers Total demand for tradables in each country is modeled as exogenous, while its composition between domestic and foreign good depend on a preference parameter. In particular, domestic consumers' problem is

$$\begin{aligned} &max\ C_h = (a_h^{\alpha} b_h^{1-\alpha}) \\ &st\ D_h = q_h a_h + q_f b_h \end{aligned} \tag{3.4}$$

where a is the domestic good and q_h its price, while b is the foreign good and q_f its price. D_h is the total consumption expenditure and it is an exogenous random variable. Consumption is a Cobb-Douglas bundle of domestic and foreign good, where α is the share of domestic consumption in domestic good. Similarly for foreign country, $C_f = (b_f^{\alpha} b_h^{1-\alpha})$, α is the share of foreign consumption in foreign good.

Market solution The market clearing conditions are

$$\begin{aligned} Y_h &= a_h + a_f \\ Y_f &= b_h + b_f \end{aligned} \tag{3.5}$$

Equation 3.5 equates supply and demand for domestic and foreign goods. Combining the consumer problem solution with the clearing condition gives

$$\begin{aligned} q_h Y_h &= \alpha D_h + (1-\alpha) D_f \\ q_f Y_f &= \alpha D_f + (1-\alpha) D_h \end{aligned} \tag{3.6}$$

Because of the simplifying assumption on the production function, in each country nominal output equates nominal labor income.

$$\begin{aligned} W_h^{NONFIN} &= \alpha D_h + (1-\alpha) D_f \\ W_f^{NONFIN} &= \alpha D_f + (1-\alpha) D_h \end{aligned} \tag{3.7}$$

As a result, non-financial income in each country is a linear function of domestic and foreign consumption expenditure in tradables, depending on the preference parameter α. Higher is α, higher is the share of domestic consumption of domestic good and therefore lower the trade between the two countries. With no trade, $\alpha = 1$, agents' labor income depend only on domestic demand. With complete trade, $\alpha = 0.5$, agents' labor income depend equally on domestic and foreign demand.

Tradable consumption demand are exogenous random variable. In particular, for each country $k \in \{h, f\}$ they follow the process

$$D_k = \bar{D}_k + c_k M_k + \epsilon_{D_k} \tag{3.8}$$

with $\epsilon_{D_k} \sim N(0, \sigma^2_{D_k})$. M_k is a country specific risk factor, distributed $M_k \sim N(0, \sigma^2_{M_k})$. Therefore, domestic demand for domestic and foreign good depends on domestic country risk factor (and similarly for foreign demand).[10] I normalize c_k to be positive, meaning that the country risk factor positively affect non-financial income.

[10] The iid error terms ϵ_{D_k} can be normalized to zero, they do not affect the model solution in any way.

Financial income

Income from portfolio investment is standard. When shocks realize in third stage, the agent receives the payoff from his portfolio of domestic and foreign asset, plus the returns on his saving in risk-free asset.

$$W_k^{FIN} = x_h f_h + x_f f_f + (W_0 - p_h - p_f)R \tag{3.9}$$

where x_h and x_f are the portfolios of home and foreign asset chosen in the previous stage; f_h and f_f are their realized returns; p_h and p_f are their prices; R is the risk free rate and W_0 the initial wealth. Similarly for foreign agents.

Similarly to domestic and foreign tradables demand, I assume that domestic and foreign asset payoffs depend on the respective country-specific risk factor and a idiosyncratic term. For $k \in \{h, f\}$,

$$f_k = \bar{f}_k + b_k M_k + \epsilon_{fk} \tag{3.10}$$

with $\epsilon_{fk} \sim N(0, \sigma_{fk}^2)$. M_k is again the country-specific risk factor and ϵ_{fk} is a idiosyncratic term.[11]

Equations 3.7 and 3.9 shows that the two component of financial wealth are both affected by domestic and foreign risk factors M_k, respectively by tradables demands D_k and asset returns f_k. However, while the exposure of non-financial income to risk factors is exogenous and given by the preference parameter α, agents can decide how to form their financial portfolio in terms of domestic and foreign risk. I use the model to investigate how the portfolio choices respond to exogenous changes in the trade parameter α.

Throughout the paper I assume $b_k > 0$, meaning that country risk factor posi-

[11] Differently from the tradable demand shocks, here the iid term is important for result. It corresponds to the "unlearnable" component of asset return similarly to Valchev (2017).

tively affect asset return. Since I also normalize $c_k > 0$, it means that non financial and financial income are positively correlated, as first shown by Baxter and Jermann (1997). In Section 3.5 I consider the case in which the correlation is negative.

3.3.3 Second stage: investment choice

In this stage agents receive signals about the country-specific risk factors and maximize their mean variance utility by allocating their resources on domestic and foreign assets. In order to compute expected value and variance of consumption, they form posterior about the risk factors using the information available: priors, asset prices and private signals.

The problem for the domestic agent is[12]

$$\max_{x_h, x_f} \qquad U_2 = E(W|I^{(i)}) - \frac{\gamma}{2}Var(W|I^{(i)}) \tag{3.11}$$

where $W = \underbrace{x_h f_h + x_f f_f + (W_0 - p_h - p_f)R}_{W^{FIN}} + \underbrace{\alpha D_h + (1-\alpha)D_f}_{W^{NONFIN}}$

His information set contains their prior, $M_k \sim N(0, \sigma^2_{Mk})$, $k \in \{h, f\}$, the market price for each asset and two private unbiased signals about the risk factors value

$$\begin{aligned} \eta_h^{(i)} &= M_h + \epsilon_h^{(i)} \qquad & \epsilon_h \sim N(0, \sigma^2_{\eta_h}) \\ \eta_f^{(i)} &= M_f + \epsilon_f^{(i)} \qquad & \epsilon_f \sim N(0, \sigma^2_{\eta_f}) \end{aligned} \tag{3.12}$$

The agent observes the signal, forms posteriors and decides his portfolio allocation. Similarly for the foreign agent.

3.3.4 First Stage: information acquisition

In this stage agents maximize their expected mean variance utility by choosing the distribution from which to draw the two private signals in the next stage. In other words, they use their attention to decrease signals variance given the information

[12] Since I consider only domestic investors, I drop the pedix notation for simplicity.

set available at this stage, which contains only the priors.

$$
\begin{aligned}
&\max_{\sigma^2_{\eta_h},\sigma^2_{\eta_f}} \qquad U_1 = E[E(W|I^{(i)}) - \frac{\gamma}{2}Var(W|I^{(i)})] \\
st \qquad & \underbrace{\frac{1}{2}\left(ln(Var(M_h|I^p)) - ln(Var(M_h|I^{(i)}))\right)}_{\kappa_h} + \underbrace{\frac{1}{2}\left(ln(Var(M_f|I^p)) - ln(Var(M_f|I^{(i)}))\right)}_{\kappa_f} \leq \kappa \\
&\kappa_k \geq 0
\end{aligned}
\tag{3.13}
$$

The implication of the utility function in equation 3.13 is a *preference for early resolution of uncertainty*: agents want to minimize the perceived variance of final wealth.[13]

Agents choose attention allocation subject to two constraints: (i) the *capacity constraint* limits the amount of information they can learn, measured as the distance between posterior variance conditional on the private information $I^{(i)} = \{\eta_h^{(i)}, \eta_f^{(i)}, p_h, p_f\}$ and posterior variance conditional on only public signals $I^p = \{p_h, p_f\}$;[14] (ii) the *no-negative-learning constraint* rules out the possibility of increasing initial uncertainty (forgetting information). The intuition is that agents are rational but limited in their capacity of processing information. As a result, they have to decide whether to focus on domestic risk factor, foreign risk factor or a combination of both. However, they can not "forget" information they already know.

3.4 Model Solution

The model is solved backwards: first I show the optimal investment allocation for a given signal precision, then how attention choice interacts with the optimal portfo-

[13] The increasing return to information in absence of trade relies on this particular utility function form. The technical details in Van Nieuwerburgh and Veldkamp (2010).

[14] By increasing the signal precision, agents increase the posterior precision.

lio.

3.4.1 Optimal portoflio

From stage 2's investment maximization problem, the optimal portfolio of domestic agent is

$$
\begin{aligned}
x_h &= \frac{\bar{f}_h + b_h \hat{M}_h^{(i)} - p_h R}{\gamma(b_h^2 \hat{\sigma}_h^2 + \sigma_{fh}^2)} - \alpha c_h b_h \frac{\hat{\sigma}_h^2}{b_h^2 \hat{\sigma}_h^2 + \sigma_{fh}^2} \\
x_f &= \frac{\bar{f}_f + b_f \hat{M}_f^{(i)} - p_f R}{\gamma(b_f^2 \hat{\sigma}_f^2 + \sigma_{ff}^2)} - (1-\alpha) c_f b_f \frac{\hat{\sigma}_f^2}{b_f^2 \hat{\sigma}_f^2 + \sigma_{ff}^2}
\end{aligned}
\tag{3.14}
$$

where $\hat{M}_k \equiv E(M_k|I^{(i)})$ is the country risk factor posterior mean and $\hat{\sigma}_k^2 \equiv Var(M_k|I^{(i)})$ is the country risk factor posterior variance.[15] Similarly for the foreign agent.

The first term in each portfolio is the Sharpe Ratio and the second term is the risk-hedging term. In particular, the Sharpe Ratio is increasing in asset return's posterior mean (numerator) and decreases in its posterior variance (denominator). The risk hedging term depends on the posterior covariance between financial and non-financial income (numerator) and asset return's posterior variance (denominator).

No information choice Consider a domestic agent and ignore first any information choice. Suppose $\alpha = 1$ (no trade): since non-financial income is positively correlated with financial income (b_h, $c_h > 0$), the agent should hedge it by investing proportionally more abroad than domestically. When $\alpha < 1$, his non-financial income becomes affected also by foreign risk, and it becomes optimal to gradually invest more domestically and less abroad. When $\alpha = 0.5$ (maximum trade), it is optimal to completely diversify the portfolio. As a result, absent any information choice, the more domestic country is trading with foreign country, the less domestic

[15] The resulting optimal allocation in the standard form $x_k = \frac{E(f_k|I) - p_k R}{\gamma Var(f_k|I)} - \frac{Cov(f_k, z|I)}{Var(f_k|I)}$ where z is a source of income correlated with the asset return.

agent should invest abroad. This theoretical results is in sharp contrast with the evidence of Section 3.2.

With information choice The previous result relies on the assumption that posterior variance of domestic and foreign country risk factor are equal. This will not be the case anymore when one allows for information choice. In particular, lower is the domestic posterior variance $\hat{\sigma}_h^2$, higher is the Sharpe Ratio of domestic asset and lower the risk-hedging term. As a result, for the same level of trade, the resulting proportional holding of domestic asset increases.

Importantly, the decrease in the risk heding term depends on the presence of the "unlearnable" risk component σ_{fk}^2 in the asset return, that makes the posterior hedging power of the asset lower when information increases.[16] If in equilibrium information increases with trade, as I show in the next sections, then the role of risk-hedging in determining portfolio allocation is downsized.

3.4.2 Equilibrium Price

In order to preserve private information in equilibrium, I make the standard assumption that in each market the supply of asset consists in a constant term plus a noisy component (noise traders) $z_k \in N(0, \sigma_z^2)$. As a result, the equilibrium market conditions are

$$\overline{z}_k + z_k = \int^H x_k^{(i)} di + \int^F x_k^{*(i)} di \qquad k \in \{h, f\} \tag{3.15}$$

The resulting price depends on the posterior distribution of the country risk factors, which in turn depend on prices as unbiased signals of country risk factor realizations. I solved this fixed point problem with a guess and verify technique:

[16] The role fo the "unlearnable" risk component has been explored by Valchev (2017).

the appendix shows that a linear solution of this problem has the form[17]

$$p_k = \overline{\lambda}_k + \lambda_{M_k} M_k + \lambda_{z_k} z_k \tag{3.16}$$

As a result, price $\tilde{p}_k$ is an unbiased signal of the country risk factor M_k

$$\frac{\tilde{p}_k - \overline{\lambda}_k}{\lambda_{M_k}} = M_k + \frac{\lambda_{z_k}}{\lambda_{M_k}} z_k \tag{3.17}$$

Agents use public (price) and private (priors and signals) information to form posteriors about country risk factors' variance and mean:

$$\begin{aligned} \hat{\sigma}_k^2 &\equiv Var(M_k|I^{(i)}) = \left(\frac{1}{\sigma_k^2} + \frac{\lambda_{Mk}^2}{\lambda_{z_k}^2 \sigma_k^2} + \frac{1}{\sigma_{\eta k}^2} \right)^{-1} \\ \hat{M}_k &\equiv E(M_k|I^{(i)}) = \hat{\sigma}_k^2 \left(\frac{\lambda_{Mk}^2}{\lambda_{z_k}^2 \sigma_k^2} \tilde{p}_k + \frac{1}{\sigma_{\eta k}^2} \eta_k^{(i)} \right) \end{aligned} \tag{3.18}$$

3.4.3 Attention allocation

Substituting for the optimal portfolio allocation and taking expectation conditional on first stage information set, the domestic agent's attention allocation problem becomes

[17] Note that, differently from traditional information cost paper (Adamanti, 1985) the price is not a function of asset return, but country risk factor.

$$
\begin{aligned}
\max_{\sigma^2_{\eta_h},\sigma^2_{\eta_f}} E^1(U_2) = \quad & \frac{B_h^2 + A_h}{2\gamma(b_h^2\hat{\sigma}_h^2 + \sigma_{fh}^2)} - \frac{b_h^2\hat{\sigma}_h^2}{2\gamma(b_h^2\hat{\sigma}_h^2 + \sigma_{fh}^2)} \\
& + \frac{B_f^2 + A_f}{2\gamma(b_f^2\hat{\sigma}_f^2 + \sigma_{ff}^2)} - \frac{b_f^2\hat{\sigma}_f^2}{2\gamma(b_f^2\hat{\sigma}_f^2 + \sigma_{ff}^2)} \\
& - \alpha\left[b_h c_h \frac{\hat{\sigma}_h^2}{b_h^2\hat{\sigma}_h^2 + \sigma_{fh}^2} B_h + \frac{\gamma}{2}(c_h)^2 \frac{\hat{\sigma}_h^2\sigma_{fh}^2}{b_h^2\hat{\sigma}_h^2 + \sigma_{fh}^2}\right] \\
& - (1-\alpha)\left[b_f c_f \frac{\hat{\sigma}_f^2}{b_f^2\hat{\sigma}_f^2 + \sigma_{ff}^2} B_f + \frac{\gamma}{2}(c_f)^2 \frac{\hat{\sigma}_f^2\sigma_{ff}^2}{b_f^2\hat{\sigma}_f^2 + \sigma_{ff}^2}\right] \\
& - \frac{\gamma}{2}^2 \sigma_y^2 \left[\alpha^2 + (1-\alpha)^2\right] + \bar{Y}_h \\
st \qquad & \kappa_h + \kappa_f \le \kappa \\
& \kappa_k \ge 0 \qquad k \in \{h, f\}
\end{aligned}
\tag{3.19}
$$

The terms B_k and A_k are functions of priors (assumed equal across countries), prices and other terms taken as given by agents. They are defined in Appendix B.

The first two rows of (3.19) derive from the risk-adjusted posterior mean return of, respectively, domestic and foreign asset (they do not depend on the hedging term). Utility in the first (second) row decreases in home (foreign) country risk posterior variance, and therefore increases in attention to it.

The third and fourth rows of (3.19) derive from the hedging term of the optimal portfolio (first term) and from the volatility of trade consumption (second term). Under the assumption of positive covariance between country's good demand and asset return ($b_k, c_k > 0$), utility in the third (fourth) row decreases in home (foreign) country risk posterior variance, and therefore increases attention to it. In the case of negative covariance ($b_k > 0, c_k < 0$ or $b_k < 0, c_k > 0$), the opposite is true.

The fifth row of (3.19) derives from tradables consumption variance caused by the endowment idiosyncratic shock plus a constant, and it does not affected by information choice.

The interaction between trade and investment makes the information problem less convex with respect to the baseline in Van Nieuwerburgh and Veldkamp (2009). The corner solution is only a particular case in a range of different possibilities depending on the parameter values. In the next section I explain in details this result.

3.5 Results

In the following sections I perform a comparative statics analysis and highlight the impact of trade on portfolio allocation. From now on I assume symmetry between the two countries in terms of parameters and idiosyncratic shocks.

3.5.1 Financial and non-financial income non correlated

Suppose that non-financial income is not correlated with asset return ($c_k = 0$). Hence, there is no hedging term in portfolio allocation and signals are informative only about asset payoffs. Formally, the attention allocation problem consists in only the first two lines of (3.19). The problem reduces to the standard endogenous information model, where the feedback between information and investment choice leads to increasing return to attention. As in the baseline case, the information problem is convex in posteriors and there are two corner solutions: complete attention to domestic or foreign asset. Agents allocate attention to only one asset, which is perceived with higher risk-adjust return and overweight in the optimal portfolio.

Figure 3.1 illustrates the information choice problem faced by the domestic agent. The red dotted line is the capacity constraint, the black solid lines the prior variances. Because of the no negative learning constraint, posterior variances can not be increased above the priors: the feasible choices are below the priors and above the capacity constraint. The blue solid line represent the highest achievable utility (it increases in indifference curves closer to the origin). The orange dotted line shows one of the two corner solutions. This particular case of my model is similar to Van Nieuwerburgh and Veldkamp (2009). I now show how trade leads to different implications.

3.5.2 Financial and non-financial income positively correlated

Allow now country risk factors to affect positively not only asset return but also non-financial income ($c_k, b_k > 0$). There are two implications: first, optimal portfolios present a hedging term; second, by decreasing posterior variance, attention lowers both the hedging term ad perceived non-financial income volatility. Formally, rows three and four in the information problem (3.19) are different from zero.

The following results depend on my calibration of the model, which I discuss in the next section.

No trade

First, suppose there is no trade ($\alpha = 1$). Non-financial income depends only on domestic factors (fourth row of (3.19) is zero). Without any attention choice, domestic agents should hold a larger share of portfolio in foreign asset, in order to hedge non-financial risk (as in Baxter and Jermann (1997)). However, they also have incentive to specialize attention in domestic risk to decrease non financial income uncertainty. Because of the increasing return to information, they end up allocating all attention to domestic asset, which becomes perceived as less risky. As a result, contrary to the risk hedging prediction, they end up overweight their own asset in their portfolio. There are two forces at work to offset the risk hedging: (i) the lower posterior risk makes the asset more attractive; (ii) the lower posterior covariance makes the asset less suitable for risk-hedging.

Figure 3.2 illustrate the information choice problem faced by the domestic agent. The indifference curve has higher slope than in the standard case, and the optimal choice is to decrease domestic posterior variance by allocating all attention to "domestic" signal and no attention to "foreign" signal.[18]

Maximum trade

Now suppose there is maximum trade ($\alpha = 0.5$). Non-financial income is equally exposed to domestic and foreign factor and agents hedge against both. Without information choice, portfolios should be perfectly diversified. Information choice does not change this result: domestic agent has no incentive to pay more attention to the domestic asset than to the foreign asset, since they are equally exposed to both.[19] Formally, the indifference curve is symmetric. Figure 3.3 illustrate this problem.

The intuition is the following: learning problem's convexity in the baseline model derives from the possibility of contemporaneously decreasing both attention and investment to one asset while focusing on the other. As a result, we have a portfolio specialization on one country risk factor. Here, by imposing final wealth to be exposed to both risk factors through trade, the increasing return to information is considerably weakened.

General case

While for $\alpha = 1$ (no trade) the model predicts home bias in investment and for $\alpha = 0.5$ (maximum trade) it predicts full diversification, for values $\alpha \in (0.5, 1)$ we have gradual opening of portfolio and attention to foreign equities: higher the trade openness, higher the portfolio diversification.

This result in shown in Figure 3.4. The horizontal axis measures the trade parameter α and the vertical axis the posterior variance of the two assets. Higher the α, lower the trade, higher the specialization of attention in domestic equities. A similar pattern is shown is Figure 3.5, that relates trade and portfolio holdings:

[18] Figure 3.2 does not show an exact corner solution, it depends on the calibration.
[19] This is a necessary but not sufficient condition to get this result. The size of the non-learnable risk component of asset payoff plays a crucial role in weaking the increasing return to information.

higher the α, lower the trade, higher the specialization on home equities.

Figure 3.4 and 3.5 show the main result of the paper: information choice creates a positive link between trade in goods and equity investment. The model can explain both the lack of diversification in presence of trade and the equity home bias when trade is low. In the latter case, it does not relies on heterogeneous priors and it reverses the risk hedging diversification prediction.

Appendix C discussed model's results with alternative parametrizations.

3.6 Trade, investment and attention in the data

The main implication of my model is that international trade affects portfolio diversification by increasing attention allocation to foreign country. I test this prediction in the data.

In order to study the impact of trade on information choice, I need to measure attention allocation. However, attention is not directly observable. I follow the financial literature (Da et al., 2011; Andrei and Hasler, 2014) in using the volume of research queries from GoogleTrend as a proxy.[20] I measure the attention allocation of country H to country F financial assets with the index of Google research frequencies in country H with the keyword ”F” in the category ”finance”.[21] The index has values 0-100, where 100 is the highest value in the sample downloaded.[22] I only have a limited sample of around 30 source and 110 destination countries, from 2005 to 2015.[23] I aggregate the monthly data to an annual frequency, and use it as a dependent variable in the following regression.

[20] Mondria et al. (2010) measure the international investors' attention allocation in a similar way, but only for the US and with a slightly different proxy.

[21] For example, the attention allocated by Italy to Germany is measured as the number of researches on Google in Italy with the keyword ”Germany” in the finance section, relative to the total number of researches in Italy.

[22] To keep the normalization consistent across observations, i always use the same highest normalizing observation.

[23] It is not possible to download them all in one time, I could only get the data on this countries.

$$ln(Att_{ijt}) = \phi_{it} + \phi_{jt} + \beta_1 ln(Trade_{ijt}) + \beta_2 ln(Equity_{ijt}) + \beta_3 X_{ijt} + \epsilon_{ijt} \quad (3.20)$$

The dependent variable is the logarithm of the volume research index, while the independent variables are (i) trade, measured as export plus import; (ii) the amount of country j equity in country i household's portfolio, and (iii) a the bilateral controls: cultural linkages (common language, colonial relationship, common religion), distance between capitals, common monetary union, GDP correlation. I include again source and destination-time fixed effect, which captures every country-specific determinants of attention. I consider only households' portfolio instead of total holding for two reason: first, to address the concern that the information set of trading firms might have other determinants, as explained in Section 3.2; (ii) because the google search volume index is typically indicated as a proxy for attention allocation of retail investors, not big financial companies.

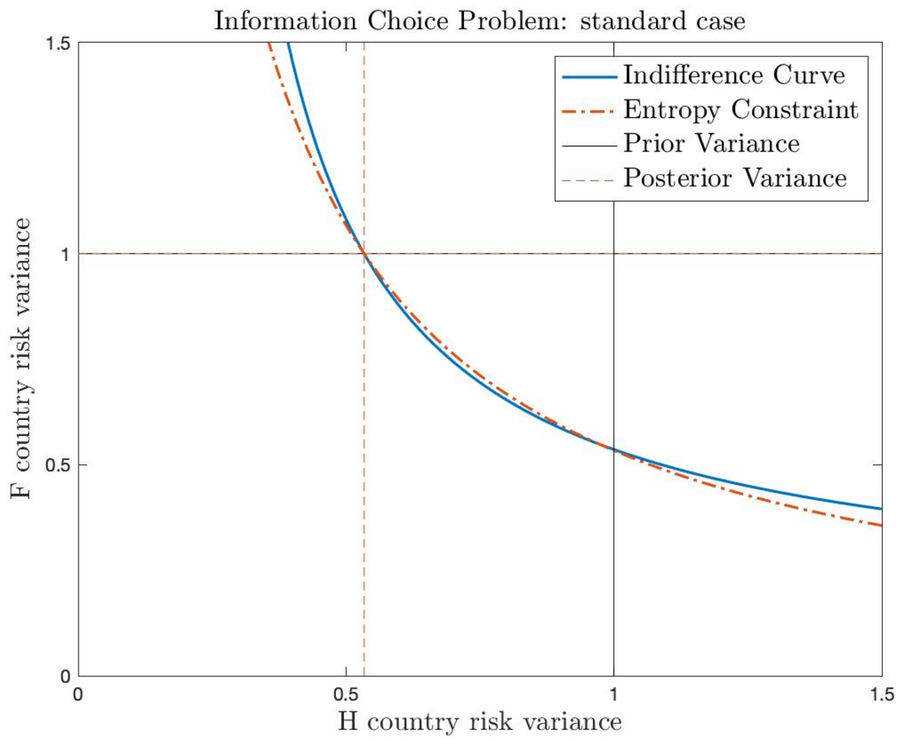
The hypothesis I am testing is that, for a given amount of equity, higher trade between countries leads to higher attention allocated: $\beta_1 > 0$. Table 3.5 confirm this prediction. Column (1) show that simple OLS coefficient for trade is positive and significant: even controlling for portoflio investment, higher trade with the foreign country increases the attention allocated to its financial characteristic. Column (2) show that this result is robust when instrumenting trade similarly to Section 3.2 while columns (3) and (4) show that the results hold when using total asset instead of only equity. Table 3.6 consider only export instead of total trade, but the result is unaffected.

3.7 Conclusions

This paper provides an answer to a common puzzle in the international economic literature: the empirical positive correlation between trade in goods and portfolio investment across countries. I present a theoretical model to rationalize the impact of trade on equity investment through information choice, and provide empirical support for my model's predictions in the data.

The model extends the baseline framework of the endogenous information literature to include income from trade. I show how the model can rationalize the positive relation between trade and portfolio diversification: trading countries optimally chose to acquire more information about each other with respect to less trading countries, thus perceiving partner's asset as less risky and more profitable. In general, higher the trade between countries, higher is their bilateral portfolio investment.

Figure 3.1: Information Stage maximization problem: standard case ($c = 0$)



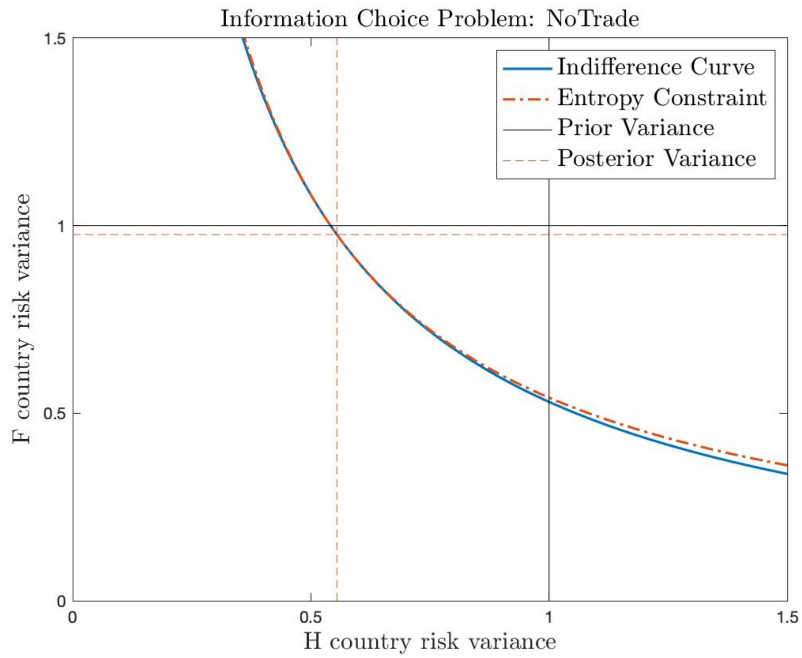


Figure 3.2: Information Stage maximization problem: no trade case ($\alpha = 1$)

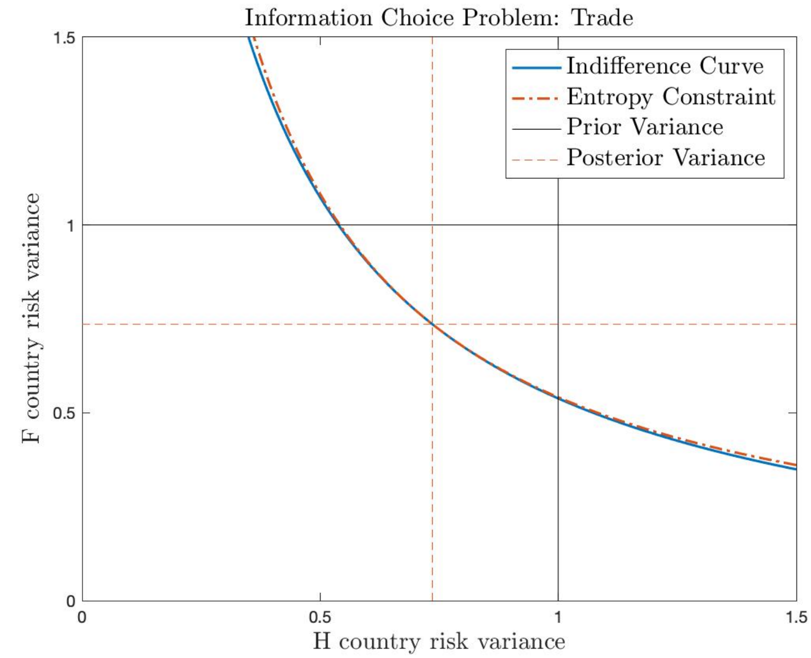


Figure 3.3: Information Stage maximization problem: trade case ($\alpha = 0.5$)

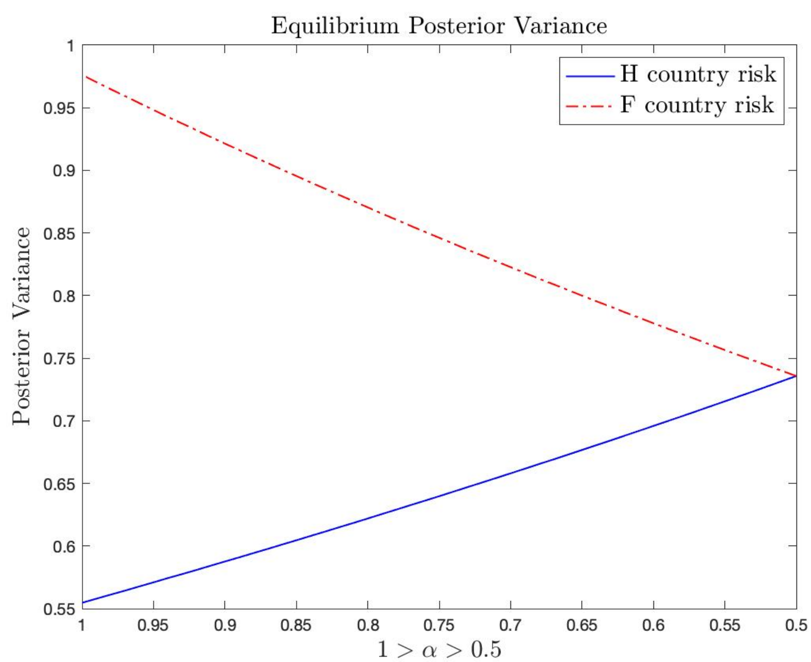


Figure 3.4: Equilibrium posterior variance for values $\alpha \in [0.5, 1]$

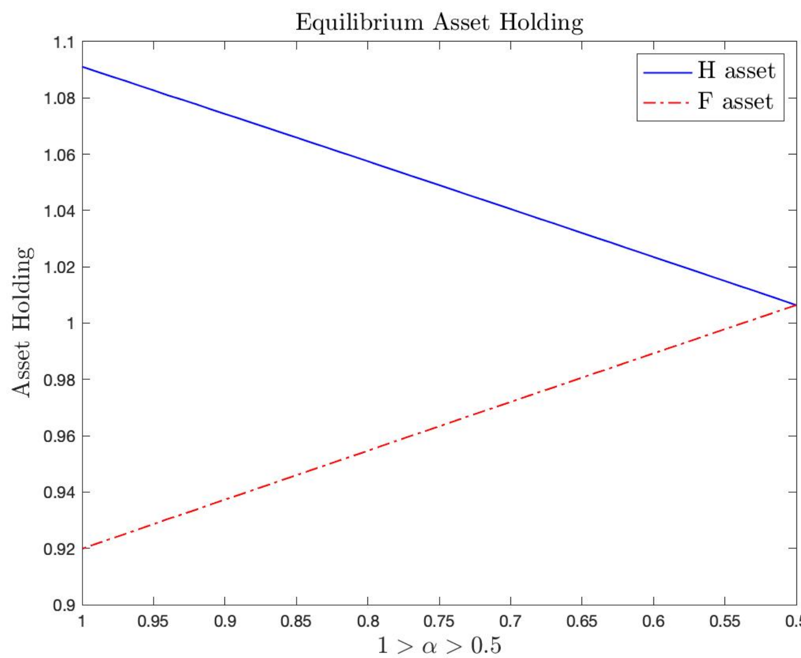


Figure 3.5: Equilibrium portfolio allocation for values $\alpha \in [0.5, 1]$

Table 3.1: Impact of trade on portoflio (total holders)

	(1) logEquity OLS	(2) logEquity IV	(3) logTotinv OLS	(4) logTotinv IV
logDistance	-0.232***	-0.392***	-0.0967	-0.404***
	(0.0735)	(0.0906)	(0.0603)	(0.0717)
logTrade	0.760***	0.677***	0.751***	0.523***
	(0.0438)	(0.0729)	(0.0313)	(0.0500)
CommonLanguage	0.311*	0.463***	0.296**	0.455***
	(0.165)	(0.161)	(0.135)	(0.134)
Colony	0.0848	0.0384	0.0188	0.0589
	(0.164)	(0.164)	(0.153)	(0.145)
CommonLegalSystem	0.472***	0.407***	0.298***	0.293***
	(0.105)	(0.103)	(0.0844)	(0.0855)
CommunCurrency	0.718***	0.466***	1.576***	1.151***
	(0.151)	(0.153)	(0.126)	(0.132)
CorrGDP	-0.438**	-0.493**	0.402**	0.490***
	(0.200)	(0.202)	(0.165)	(0.169)
CommonBorder	0.0292	0.142	-0.342**	-0.179
	(0.177)	(0.173)	(0.162)	(0.165)
R-squared	0.735	0.171	0.729	0.211
N	18863	17127	25630	23182
Weakid F		229		374
Underid p		0.000		0.000
Hansen p		0.244		0.087

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3.2: Impact of export on portoflio (total holders)

	(1) logEquity OLS	(2) logEquity IV	(3) logTotinv OLS	(4) logTotinv IV
logDistance	-0.360***	-0.463***	-0.194***	-0.433***
	(0.0718)	(0.0866)	(0.0601)	(0.0715)
logExport	0.602***	0.566***	0.617***	0.457***
	(0.0376)	(0.0639)	(0.0288)	(0.0461)
CommonLanguage	0.253	0.453***	0.228*	0.435***
	(0.166)	(0.161)	(0.136)	(0.134)
Colony	0.155	0.0805	0.0533	0.0646
	(0.170)	(0.164)	(0.157)	(0.146)
CommonLegalSystem	0.525***	0.429***	0.361***	0.311***
	(0.107)	(0.103)	(0.0862)	(0.0862)
CommunCurrency	0.688***	0.442***	1.554***	1.132***
	(0.153)	(0.150)	(0.129)	(0.131)
CorrGDP	-0.327	-0.438**	0.466***	0.508***
	(0.199)	(0.201)	(0.163)	(0.167)
CommonBorder	0.0349	0.141	-0.322*	-0.170
	(0.182)	(0.176)	(0.166)	(0.166)
R-squared	0.729	0.161	0.723	0.202
N	18895	17147	25696	23216
Weakid F		207		343
Underid p		0.000		0.000
Hansen p		0.194		0.092

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3.3: Impact of trade on portoflio (HH holders)

	(1) logEquity OLS	(2) logEquity IV	(3) logTotinv OLS	(4) logTotinv IV
logDistance	-0.489***	-0.751***	-0.489***	-0.648***
	(0.137)	(0.136)	(0.137)	(0.152)
logTrade	0.655***	0.319***	0.655***	0.540***
	(0.0716)	(0.107)	(0.0716)	(0.133)
CommonLanguage	0.273	0.202	0.273	0.516*
	(0.318)	(0.247)	(0.318)	(0.285)
Colony	0.687**	0.794***	0.687**	0.482*
	(0.267)	(0.273)	(0.267)	(0.265)
CommonLegalSystem	0.863***	0.770***	0.863***	0.806***
	(0.155)	(0.137)	(0.155)	(0.158)
CommunCurrency	1.391***	0.720***	1.391***	0.958***
	(0.231)	(0.196)	(0.231)	(0.216)
CorrGDP	-0.635*	-0.781**	-0.635*	-1.001***
	(0.334)	(0.325)	(0.334)	(0.344)
CommonBorder	-0.163	0.459**	-0.163	0.156
	(0.274)	(0.212)	(0.274)	(0.246)
R-squared	0.707	0.101	0.707	0.097
N	6766	5161	6766	5964
Weakid F		118		90
Underid p		0.000		0.000
Hansen p		0.776		0.040

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3.4: Impact of export on portoflio (HH holders)

	(1) logEquity OLS	(2) logEquity IV	(3) logTotinv OLS	(4) logTotinv IV
logDistance	-0.621***	-0.798***	-0.621***	-0.764***
	(0.137)	(0.125)	(0.137)	(0.130)
logExport	0.527***	0.279***	0.527***	0.408***
	(0.0633)	(0.0932)	(0.0633)	(0.0942)
CommonLanguage	0.218	0.190	0.218	0.498*
	(0.322)	(0.247)	(0.322)	(0.286)
Colony	0.798***	0.825***	0.798***	0.571**
	(0.273)	(0.272)	(0.273)	(0.262)
CommonLegalSystem	0.901***	0.785***	0.901***	0.833***
	(0.159)	(0.137)	(0.159)	(0.157)
CommunCurrency	1.380***	0.737***	1.380***	0.920***
	(0.235)	(0.198)	(0.235)	(0.211)
CorrGDP	-0.565*	-0.808**	-0.565*	-0.908***
	(0.331)	(0.325)	(0.331)	(0.328)
CommonBorder	-0.182	0.447**	-0.182	0.172
	(0.282)	(0.214)	(0.282)	(0.251)
R-squared	0.703	0.094	0.703	0.101
N	6776	5164	6776	5968
Weakid F		68		114
Underid p		0.000		0.000
Hansen p		0.662		0.036

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3.5: Impact of trade on attention (HH holders)

	(1) LogAttention OLS	(2) LogAttention IV	(3) LogAttention OLS	(4) LogAttention IV
logDistance	-0.175***	-0.114*	-0.109**	-0.0690
	(0.0549)	(0.0621)	(0.0425)	(0.0531)
logTrade	0.123***	0.282***	0.108***	0.266***
	(0.0194)	(0.0349)	(0.0170)	(0.0353)
LogEquity	0.0216***	0.0230***		
	(0.00824)	(0.00758)		
logAsset			0.0327***	0.0326***
			(0.00819)	(0.00843)
CommonLanguage	0.437***	0.347***	0.605***	0.461***
	(0.109)	(0.0980)	(0.102)	(0.0872)
Colony	0.204**	0.147	0.112	0.0717
	(0.103)	(0.104)	(0.0914)	(0.0921)
CommonLegalSystem	0.0202	-0.0264	0.0270	-0.0655
	(0.0454)	(0.0464)	(0.0396)	(0.0426)
CommunCurrency	-0.169**	-0.196***	-0.211***	-0.224***
	(0.0690)	(0.0614)	(0.0649)	(0.0605)
CorrGDP	0.229*	0.230*	0.00653	0.164*
	(0.125)	(0.125)	(0.0989)	(0.0921)
CommonBorder	-0.107	-0.204***	0.0136	-0.139*
	(0.0797)	(0.0763)	(0.0770)	(0.0747)
CommonRelig	0.00603	-0.00496	0.0853	0.110
	(0.0838)	(0.0752)	(0.0740)	(0.0677)
R2	0.654	0.124	0.625	0.153
N	5535	4328	6957	5319
Weakid F		296		77
Underid p		0.000		0.000
Hansen p		0.291		0.674

Standard errors in parentheses
Errors clustered at the country-pair level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Table 3.6: Impact of export on attention (HH holders)

	(1) LogAttention OLS	(2) LogAttention IV	(3) LogAttention OLS	(4) LogAttention IV
logDistance	-0.160***	-0.128**	-0.102**	-0.0751
	(0.0534)	(0.0593)	(0.0418)	(0.0524)
logExport	0.139***	0.262***	0.116***	0.251***
	(0.0179)	(0.0343)	(0.0163)	(0.0349)
LogEquity	0.0204***	0.0247***		
	(0.00790)	(0.00742)		
logAsset			0.0314***	0.0301***
			(0.00782)	(0.00832)
CommonLanguage	0.427***	0.334***	0.600***	0.459***
	(0.106)	(0.0948)	(0.102)	(0.0874)
Colony	0.214**	0.171*	0.115	0.0859
	(0.101)	(0.103)	(0.0908)	(0.0925)
CommonLegalSystem	0.0165	-0.0172	0.0286	-0.0543
	(0.0449)	(0.0453)	(0.0391)	(0.0416)
CommunCurrency	-0.159**	-0.183***	-0.207***	-0.223***
	(0.0686)	(0.0605)	(0.0648)	(0.0609)
CorrGDP	0.211*	0.194	0.00655	0.171*
	(0.123)	(0.123)	(0.0996)	(0.0934)
CommonBorder	-0.116	-0.210***	0.00267	-0.147**
	(0.0779)	(0.0721)	(0.0757)	(0.0717)
CommonRelig	0.00786	0.00461	0.0917	0.131*
	(0.0819)	(0.0739)	(0.0729)	(0.0671)
R2	0.660	0.139	0.628	0.148
N	5537	4329	6964	5321
Weakid F		118		66
Underid p		0.000		0.000
Hansen p		0.537		0.990

Standard errors in parentheses
Errors clustered at the country-pair level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.010$

Bibliography

ANDREI, D. AND M. HASLER (2014): “Investor attention and stock market volatility,” *The review of financial studies*, 28, 33–72.

ANGELETOS, G.-M. AND C. LIAN (2017): “Dampening general equilibrium: From micro to macro,” Tech. rep., National Bureau of Economic Research.

AVIAT, A. AND N. COEURDACIER (2007): “The geography of trade in goods and asset holdings,” *Journal of International Economics*, 71, 22–51.

BARON, M. AND W. XIONG (2017): “Credit expansion and neglected crash risk,” *The Quarterly Journal of Economics*, 132, 713–764.

BASU, S. (2019): “Are price-cost markups rising in the united states? a discussion of the evidence,” *Journal of Economic Perspectives*, 33, 3–22.

BAXTER, M. AND U. J. JERMANN (1997): “The international diversification puzzle is worse than you think,” *The American Economic Review*, 87, 170.

BEBCHUK, L. A., A. COHEN, AND H. SPAMANN (2010): “The wages of failure: Executive compensation at Bear Stearns and Lehman 2000-2008,” *Yale J. on Reg.*, 27, 257.

BENIGNO, P. AND S. NISTICO (2012): “International portfolio allocation under model uncertainty,” *American economic Journal: macroeconomics*, 4, 144–89.

BIANCHI, J. AND E. G. MENDOZA (2020): "A fisherian approach to financial crises: Lessons from the sudden stops literature," *Review of Economic Dynamics*, 37, S254–S283.

BLINDER, A. S. (2009): "Crazy Compensation and the Crisis," *Wall Street Journal*.

BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2018a): "Over-reaction in macroeconomic expectations," Tech. rep., National Bureau of Economic Research.

——— (2020): "Overreaction in macroeconomic expectations," *American Economic Review*, 110, 2748–82.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2018b): "Diagnostic expectations and credit cycles," *The Journal of Finance*, 73, 199–227.

BORDALO, P., N. GENNAIOLI, A. SHLEIFER, AND S. TERRY (2019): "Real Credit Cycles," Tech. rep., Mimeo, Harvard University.

BORDALO, P., N. GENNAIOLI, A. SHLEIFER, AND S. J. TERRY (2021): "Real credit cycles," Tech. rep., National Bureau of Economic Research.

BOTTAZZI, L., P. PESENTI, AND E. VAN WINCOOP (1996): "Wages, profits and the international portfolio puzzle," *European Economic Review*, 40, 219–254.

BRANCH, B. (2002): "The costs of bankruptcy: A review," *International Review of Financial Analysis*, 11, 39–57.

BRENNAN, M. J. AND H. H. CAO (1997): "International portfolio investment flows," *The Journal of Finance*, 52, 1851–1880.

BRIS, A., I. WELCH, AND N. ZHU (2006): "The costs of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization," *The journal of finance*, 61, 1253–1303.

BROER, T. AND A. KOHLHAS (2018): "Forecaster (mis-) behavior," .

BRUNNERMEIER, M. K. AND Y. SANNIKOV (2014): "A macroeconomic model with a financial sector," *American Economic Review*, 104, 379–421.

CASTRO, R., G. L. CLEMENTI, AND Y. LEE (2015): "Cross sectoral variation in the volatility of plant level idiosyncratic shocks," *The Journal of Industrial Economics*, 63, 1–29.

CBO (2017): "International comparisons of corporate income tax rates," Tech. rep., Exhibit 7.

CHODOROW-REICH, G. (2014): "The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis," *The Quarterly Journal of Economics*, 129, 1–59.

COEURDACIER, N. (2009): "Do trade costs in goods market lead to home bias in equities?" *Journal of international Economics*, 77, 86–100.

COEURDACIER, N. AND P.-O. GOURINCHAS (2016): "When bonds matter: Home bias in goods and assets," *Journal of Monetary Economics*, 82, 119–137.

COEURDACIER, N., H. REY, AND P. WINANT (2011): "The risky steady state," *American Economic Review*, 101, 398–401.

COIBION, O. AND Y. GORODNICHENKO (2012): "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy*, 120, 116–159.

——— (2015): "Information rigidity and the expectations formation process: A simple framework and new facts," *American Economic Review*, 105, 2644–78.

COIBION, O., Y. GORODNICHENKO, AND S. KUMAR (2018): "How do firms form their expectations? new survey evidence," *American Economic Review*, 108, 2671–2713.

Cole, S., M. Kanz, and L. Klapper (2014): "Incentivizing Calculated Risk-Taking: Evidence from an Experiment with Commercial Bank Loan Officers," *Journal of Finance*.

Croushore, D. D. (1997): "The Livingston survey: Still useful after all these years," *Business Review-Federal Reserve Bank of Philadelphia*, 2, 1.

Crowe, C. (2010): "Consensus forecasts and inefficient information aggregation," *IMF Working Papers*, 1–43.

Da, Z., J. Engelberg, and P. Gao (2011): "In search of attention," *The Journal of Finance*, 66, 1461–1499.

Dasgupta, K. and J. Mondria (2018): "Inattentive importers," *Journal of International Economics*, 112, 150–165.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 135, 561–644.

Dell'Ariccia, G., L. Laeven, and R. Marquez (2014): "Real interest rates, leverage, and bank risk-taking," *Journal of Economic Theory*, 149, 65–99.

Dupor, B., T. Kitamura, and T. Tsuruga (2010): "Integrating Sticky Information and Sticky Prices," *Review of Economics and Statistics*, 92, 657–669.

Durrant, J., J. J. Gong, and J. Howard (2020): "In the Nick of Time: Performance-Based Compensation and Proactive Responses to the Tax Cuts and Jobs Act," *Journal of Management Accounting Research*.

Edmans, A., X. Gabaix, and D. Jenter (2017): "Executive compensation: A survey of theory and evidence," in *The handbook of the economics of corporate governance*, Elsevier, vol. 1, 383–539.

European Central Bank (2014): "Results of the Second Special Questionnaire for Participants in the ECB Survey of Professional Forecasters," .

Farhi, E. and I. Werning (2020): "Taming a Minsky Cycle," Tech. rep., Working Paper.

Galstyan, V., P. R. Lane, C. Mehigan, and R. Mercado (2016): "The holders and issuers of international portfolio securities," *Journal of the Japanese and International Economies*, 42, 100–108.

Gehrig, T. (1993): "Intermediation in search markets," *Journal of Economics & Management Strategy*, 2, 97–120.

Gemmi, L. and R. Valchev (2021): "Public Information in Expectation Surveys," .

Gennaioli, N. and A. Shleifer (2010): "What comes to mind," *The Quarterly journal of economics*, 125, 1399–1433.

——— (2018): *A crisis of beliefs: Investor psychology and financial fragility*, Princeton University Press.

Gertler, M. and N. Kiyotaki (2010): "Financial intermediation and credit policy in business cycle analysis," in *Handbook of monetary economics*, Elsevier, vol. 3, 547–599.

Greenwood, R., S. G. Hanson, and L. J. Jin (2019): "Reflexivity in credit markets," Tech. rep., National Bureau of Economic Research.

Greenwood, R., S. G. Hanson, A. Shleifer, and J. A. Sørensen (2020): "Predictable financial crises," Tech. rep., National Bureau of Economic Research.

Gulen, H., M. Ion, and S. Rossi (2019): "Credit cycles, expectations, and corporate investment," *Expectations, and Corporate Investment (April 9, 2019)*.

He, Z. and A. Krishnamurthy (2013): "Intermediary asset pricing," *American Economic Review*, 103, 732–70.

——— (2019): "A macroeconomic framework for quantifying systemic risk," *American Economic Journal: Macroeconomics*, 11, 1–37.

Heathcote, J. and F. Perri (2013): "The international diversification puzzle is not as bad as you think," *Journal of Political Economy*, 121, 1108–1159.

Ho, P.-H., C.-W. Huang, C.-Y. Lin, and J.-F. Yen (2016): "CEO overconfidence and financial crisis: Evidence from bank lending and leverage," *Journal of Financial Economics*, 120, 194–209.

Hribar, P. and H. Yang (2016): "CEO overconfidence and management forecasting," *Contemporary Accounting Research*, 33, 204–227.

Jeanne, O. and A. Korinek (2019): "Managing credit booms and busts: A Pigouvian taxation approach," *Journal of Monetary Economics*, 107, 2–17.

Jordà, Ò., M. Schularick, and A. M. Taylor (2013): "When credit bites back," *Journal of Money, Credit and Banking*, 45, 3–28.

Julliard, C. (2002): "The international diversification puzzle is not worse than you think," .

Kiley, M. T. (2007): "A quantitative comparison of sticky-price and sticky-information models of price setting," *Journal of Money, Credit and Banking*, 39, 101–125.

Kindleberger, C. P. (1978): *Manias, panics, and crashes: a history of financial crises*, Basic books.

Klenow, P. J. and J. L. Willis (2007): "Sticky information and sticky prices," *Journal of Monetary Economics*, 54, 79–99.

KNOTEK II, E. S. (2010): "A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment," *Journal of Money, Credit and Banking*, 42, 1543–1564.

KOHLHAS, A. N. AND A. WALTHER (2020): "Asymmetric Attention," .

KORENOK, O. (2008): "Empirical comparison of sticky price and sticky information models," *Journal of Macroeconomics*, 30, 906–927.

KOZENIAUSKAS, N., A. ORLIK, AND L. VELDKAMP (2018): "What are uncertainty shocks?" *Journal of Monetary Economics*, 100, 1–15.

KRISHNAMURTHY, A. AND W. LI (2021): "Dissecting Mechanisms of Financial Crises: Intermediation and Sentiment," .

KRISHNAMURTHY, A. AND T. MUIR (2017): "How credit cycles across a financial crisis," Tech. rep., National Bureau of Economic Research.

LANE, P. R. (2000): "International investment positions: a cross-sectional analysis," *Journal of international money and finance*, 19, 513–534.

LANE, P. R. AND G. M. MILESI-FERRETTI (2008): "International investment patterns," *The Review of Economics and Statistics*, 90, 538–549.

LINDBECK, A. AND J. WEIBULL (2017): "Investment, rational inattention, and delegation," Tech. rep., IFN Working Paper.

LÓPEZ-SALIDO, D., J. C. STEIN, AND E. ZAKRAJŠEK (2017): "Credit-market sentiment and the business cycle," *The Quarterly Journal of Economics*, 132, 1373–1426.

MACKOWIAK, B., F. MATEJKA, AND M. WIEDERHOLT (2018): "Rational inattention: A disciplined behavioral model," Tech. rep., mimeo.

MACKOWIAK, B. AND M. WIEDERHOLT (2012): "Information processing and limited liability," *American Economic Review*, 102, 30–34.

MAĆKOWIAK, B. AND M. WIEDERHOLT (2015): "Business cycle dynamics under rational inattention," *The Review of Economic Studies*, 82, 1502–1532.

MALMENDIER, U. AND G. TATE (2005): "CEO overconfidence and corporate investment," *The journal of finance*, 60, 2661–2700.

MANKIW, N. G. AND R. REIS (2002): "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve," *The Quarterly Journal of Economics*, 117, 1295–1328.

MARINOVIC, I., M. OTTAVIANI, AND P. SORENSEN (2013): "Forecasters' objectives and strategies," in *Handbook of economic forecasting*, Elsevier, vol. 2, 690–720.

MASSA, M. AND A. SIMONOV (2006): "Hedging, familiarity and portfolio choice," *The Review of Financial Studies*, 19, 633–685.

MAXTED, P. (2019): "A macro-finance model with sentiment," *Unpublished working paper*.

MERTON, R. C. (1987): "A simple model of capital market equilibrium with incomplete information," *The journal of finance*, 42, 483–510.

MIAN, A., A. SUFI, AND E. VERNER (2017): "Household debt and business cycles worldwide," *The Quarterly Journal of Economics*, 132, 1755–1817.

MINSKY, H. P. (1977): "The financial instability hypothesis: An interpretation of Keynes and an alternative to "standard" theory," *Challenge*, 20, 20–27.

MONDRIA, J., T. WU, AND Y. ZHANG (2010): "The determinants of international investment and attention allocation: Using internet search query data," *Journal of International Economics*, 82, 85–95.

MORRIS, S. AND H. S. SHIN (2002): "Social value of public information," *american economic review*, 92, 1521–1534.

MORSE, A. AND A. VISSING-JORGENSEN (2020): "Information Transmission from the Federal Reserve to the Stock Market: Evidence from Governors' Calendars," Tech. rep., working paper.

OBSTFELD, M. AND K. ROGOFF (2000): "The six major puzzles in international macroeconomics: is there a common cause?" *NBER macroeconomics annual*, 15, 339–390.

OECD (2020): "COVID-19 Government Financing Support Programmes for Businesses," Tech. rep., OECD Paris.

OTTAVIANI, M. AND P. N. SØRENSEN (2006): "The strategy of professional forecasting," *Journal of Financial Economics*, 81, 441–466.

OTTO, C. A. (2014): "CEO optimism and incentive compensation," *Journal of Financial Economics*, 114, 366–404.

PORTES, R. AND H. REY (2005): "The determinants of cross-border equity flows," *Journal of international Economics*, 65, 269–296.

SCHULARICK, M. AND A. M. TAYLOR (2012): "Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008," *American Economic Review*, 102, 1029–61.

SHANNON, C. E. (1948): "A mathematical theory of communication," *Bell Syst. Tech. J.*, 27, 623–656.

SIMS, C. A. (2003): "Implications of rational inattention," *Journal of monetary Economics*, 50, 665–690.

——— (2006): "Rational inattention: Beyond the linear-quadratic case," *American Economic Review*, 96, 158–163.

STREBULAEV, I. A. AND T. M. WHITED (2011): "Dynamic models and structural estimation in corporate finance," *Foundations and Trends in Finance*, 6.

VALCHEV, R. (2017): "Information Acquisition and Portfolio Bias in a Dynamic World," *Available at SSRN 2787166*.

VAN NIEUWERBURGH, S. AND L. VELDKAMP (2009): "Information immobility and the home bias puzzle," *The Journal of Finance*, 64, 1187–1215.

——— (2010): "Information acquisition and under-diversification," *The Review of Economic Studies*, 77, 779–805.

VAN WINCOOP, E. AND F. E. WARNOCK (2010): "Can trade costs in goods explain home bias in assets?" *Journal of International Money and finance*, 29, 1108–1123.

WOODFORD, M. (2001): "Imperfect common knowledge and the effects of monetary policy," Tech. rep., National Bureau of Economic Research.

Appendix A

Rational Overoptimism and Moral Hazard in Credit Booms

A.1 Data

I combine data from two main sources: (1) Compustat for publicly listed US firms, (2) the Institutional Brokers Estimate System (I/B/E/S) Guidance database for manager's earnings forecast and actual earnings. In addition, I used the Execucomp database for executive compensation data and CRSP for monthly stock prices. To construct the sample, I discarded firm-year with negative values for assets and book value. Moreover, I consider only US firms reporting in US dollars and CEO compensation data.

A.1.1 Compustat and Execucomp

I downloaded the US Fundamentals Annual file in the CRSP/Compustat Merged dataset available through Wharton Research Data Services (WRDS). The variables I use are constructed from Compustat variables as follows:

- Annual return on stock: $\left(\dfrac{prcc\_f_t + dvpsx_t}{ajex_t}\right) / \left(\dfrac{prcc\_f_{t-1} + dvpsx_{t-1}}{ajex_{t-1}}\right) - 1$

- Leverage: $\dfrac{dltt}{at}$
- Market Value: $mktcap$
- Cash over assets: $\dfrac{ch}{at}$
- EBIT over assets: $\dfrac{ebit}{at}$
- Size: at
- Closing price (fiscal): $prcc_f$

I also use the CRSP database to compute the monthly stock return standard deviation as follows:

1. Monthly stock return: $\left(\dfrac{prccm_\tau + dvpsxm_\tau}{ajexm_\tau}\right) / \left(\dfrac{prccm_{\tau-1} + dvpsxm_{\tau-1}}{ajexm_{\tau-1}}\right) - 1$ with τ indicate month
2. Standard deviation of annual return in year t as the standard deviation of monthly return in last 60 months (minimum of 40 months).

From Execucomp, I considered only the current CEO from each firms and compute the following variables:

- CEO tenure: $year - becomeceo$ (drop if tenure¡0)
- Age: age
- Number of stock shares holding: $shrown\_excl\_opts$
- Value of stock shares holding: $shrown\_excl\_opts \times lprcc_f$
- Number of unexercised vested options: $opt\_unex\_exer\_num$
- Value of unexercised vested options: $opt\_unex\_exer\_est\_val$

- Dummy options: equal to 1 if number of unexercised vested options is larger than zero.
- Salary: $salary$

A.1.2 I/B/E/S

I downloaded I/B/E/S annual earnings per share (eps) forecast and realization data adjusted for stick-split from WRDS. I made the following sample restrictions

- I considered only US currency earnings ($curr = USD$), range and point forecast ($range\_desc = 01, 02$), comparable guidance ($diff\_code = 58$).
- I exclude observations where announcement dates is later than frecasted date; when firms issues multiple forecast in the same date on same horizon I keep the last forecast.

I consider forecasts released by firm i in year t about earnings of the same firm at the end of the same fiscal year. I compute forecast errors are realization minus forecasts. In order to make the errors comparable across firms, I normalize them by firm's earnings standard deviation.

- Standard deviation of realized earnings sd:
  - I first detrend each firm's earnings realization by subtracting the yearly median across firms. I use the median to lower the concerns about outliers
  - I then compute the standard deviation of individual firm's detrended earnings from 1985. I consider only firms that reports more than 10 years of data.
- Firm's forecast $E_t[eps_{it}]$: val_1 if firm provides a point forecast, and $(val_1 - val_2)/2$ if the firm provides a range forecast.

- Manager's squared forecast errors fe^2: $\left(\frac{eps_{it}-E_t[eps_{it}]}{sd}\right)^2$,

- Analysts average forecast $\tilde{E}_t[eps_{it}]$: $mean\_at\_date$

- Analysts squared forecast errors $\tilde{fe}^2$: $\left(\frac{eps_{it}-\tilde{E}_t[eps_{it}]}{sd}\right)^2$

- Forecast width: zero if firm provides a point forecast, and $(val_1 - val_2)/sd$ if the firm provides a range forecast.

- Forecast lead: difference between fiscal year end month forecasted and month the forecast is released by the firm in months.

A.1.3 Summary statistics

Variable are winsorized before the analysis and I exclude firm with less than 5 observations. I am left with a sample of around 4500 firm-year observations from 2004 to 2018.

A.2 Stage-2 equilibrium

The stage-2 equilibrium can be equivalently expressed in terms of firm's issuance of bond $\tilde{b}_j$ and bond price q_j instead of loan rate r_j and loan quantity b_j, where $q_j = \frac{1}{1+r_j}$, and $\tilde{b}_j = \frac{b_j}{q_j}$.

Information Agents observe the signal $z = \epsilon_j + \theta + \eta_j$, with $\epsilon_j \sim N(0, \sigma_\epsilon^2)$ and $\eta_j \sim N(0, \sigma_\eta^2)$, and may observe $\theta \sim N(0, \sigma_\theta^2)$. Therefore information set of agent j is $\Omega_j = \{z_j, \theta\}$ or $\Omega_j = \{z_j\}$ depending on their choice in the first stage.

Define $\tilde{z} = z - \theta$. Posteriors are $e|\tilde{z} \sim N(E[e|\tilde{z}], Var[e|\tilde{z}])$ with $E[e|\tilde{z}] = \tilde{m}\tilde{z}$ with $\tilde{m} = \frac{\sigma_e^2}{\sigma_e^2+\sigma_\eta^2}$ and $Var[e|\tilde{z}] = \frac{\sigma_e^2\sigma_\eta^2}{\sigma_e^2+\sigma_\eta^2}$, and $\theta|z \sim N(E[\theta|z], Var[\theta|z])$ with $E[\theta|z] = \delta z$ with $\delta = \frac{\sigma_\theta^2}{\sigma_e^2+\sigma_\eta^2+\sigma_\theta^2}$ and $Var[e|\tilde{z}] = \frac{\sigma_\theta^2(\sigma_e^2+\sigma_\eta^2)}{\sigma_e^2+\sigma_\eta^2+\sigma_\theta^2}$.

Table A.1: Summary statistics

	Mean	Median	Standard deviation
SquaredForError	0.16	0.02	0.72
SquaredForErrorAnalyst	0.20	0.02	0.80
ActualEPS	2.45	2.06	1.89
OptionsDummy	0.85	1.00	0.36
OptionsVal	28264.10	12902.38	43779.49
OptionsNum	665.81	376.67	911.93
Equity	39221.11	11310.77	128343.64
EquityNum	905.33	283.74	2180.37
Salary	892.97	893.75	316.59
ForLead	8.62	10.00	2.75
ForWidth	0.14	0.10	0.14
Age	56.21	56.00	6.70
Tenure	8.46	7.00	6.75
AnnualReturn	0.13	0.12	0.33
MonthlyReturnSd	0.09	0.09	0.04
MktCap	7119.60	2991.44	10584.35
BookVal	17.36	15.13	10.68
Leverage	0.20	0.20	0.14
StockPrice	44.02	39.36	24.73
TotalAssets	8090.72	2997.71	13642.88

Marginal effects

Bargaining process Define $C(\theta) = ln\left(\frac{k+\frac{1}{2}\phi k^2}{q\Lambda(M)k^\alpha}\right) - \theta$. The expected payoff of firm manager conditioning on stage-2 information set Ω_j is

$$\begin{aligned}
E[w_{firm,j}|\Omega_j] = &- \left[1 - \int_{-\infty}^{\infty}\int_{-\infty}^{C(\theta)} \phi(\epsilon_j|\theta, z_j)d\epsilon_j\phi(\theta|\Omega_j)d\theta\right]\tilde{b}_j \\
&- \left[\int_{-\infty}^{\infty}\int_{-\infty}^{C(\theta)} \phi(\epsilon_j|\theta, z_j)d\epsilon_j\phi(\theta|\Omega_j)d\theta\right]\psi c_d k_j(q_j, \tilde{b}_j) \\
&+ k_j(q_j, \tilde{b}_j)^\alpha \int_{-\infty}^{\infty}\int_{C(\theta)}^{\infty} \Lambda(\theta)e^{\epsilon_j}\phi(\epsilon_j|\theta, z_j)d\epsilon_j e^\theta\phi(\theta|\Omega_j)d\theta
\end{aligned}$$

while the expected payoff of the bank manager conditioning on stage-2 information set Ω_j is

$$\begin{aligned} E[w_{bank}|\Omega_j] =& b_j \left(\left[1 - \int_{-\infty}^{\infty} \int_{-\infty}^{C(\theta)} \phi(\epsilon_j|\theta, z_j) d\epsilon_j \phi(\theta|\Omega_j) d\theta \right] \right) \\ & - b_j \left(1 - \psi \left[\int_{-\infty}^{\infty} \int_{-\infty}^{C(\theta)} \phi(\epsilon_j|\theta, z_j) d\epsilon_j \phi(\theta|\Omega_j) d\theta \right] \right) \frac{q_j}{q} \end{aligned} \tag{A.1}$$

where $\phi(\epsilon_j|\theta, z_j) = \phi\left(\frac{C - E[\epsilon_j|\theta, z_j]}{\sqrt{Var[\epsilon_j|\theta, z_j]}}\right)$ is the posterior distribution of ϵ_j conditioning on θ and z_j, and $\phi(\theta|\Omega_j) = \phi\left(\frac{\theta - E[\theta|\Omega_j]}{\sqrt{Var[\theta|\Omega_j]}}\right)$ is the posterior distribution of θ conditioning on information set Ω_j, which may or not include θ.

Bank and firm manager decide collectively bond issued $\tilde{b}_j$ and price q_j through Nash Bargaining

$$max_{q_j, \tilde{b}_j} \; (E[w_{firm,j}|\Omega_j])^{\beta} (E[w_{bank,j}|\Omega_j])^{1-\beta} \tag{A.2}$$

Since I assume $\beta \to 1$, the problem becomes becomes

$$\begin{aligned} & max_{q_j, b_j} \; E[w_{firm,j}|\Omega_j] \\ & s.t. \quad E[w_{bank,j}|\Omega_j] \geq 0 \end{aligned} \tag{A.3}$$

Note that maximizing in terms of k_j is equivalent to maximizing in terms of $\tilde{b}_j$. The resulting first order conditions are given by

$$\begin{aligned} & E[w_{bank,j}|\Omega_j] = 0 \\ & \frac{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}}{\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial \tilde{b}_j}} = \frac{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial q_j}}{\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial q_j}} \end{aligned} \tag{A.4}$$

where each term is defined as follow. Define $pdef_j = \left[\int_{-\infty}^{\infty} \int_{-\infty}^{C(\theta)} \phi(\epsilon_j|\theta, z_j) d\epsilon_j \phi(\theta|\Omega_j) d\theta \right]$.

Then,

$$\begin{aligned}\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j} =& -[1-pdef_j] - [pdef_j]\,\psi c_d \frac{\partial k_j}{\partial \tilde{b}_j} \\ & - \left[\int_{-\infty}^{\infty} \phi(C|\theta, z_j)\frac{\partial C}{\partial \tilde{b}_j} d\epsilon_j \phi(\theta|\Omega_j)d\theta\right]\psi c_d k_j \\ & + \alpha k_j^{\alpha-1}\frac{\partial k_j}{\partial \tilde{b}_j}\int_{-\infty}^{\infty}\int_{C(\theta)}^{\infty}\Lambda(\theta)e^{\epsilon_j}\phi(\epsilon_j|\theta, z_j)d\epsilon_j e^{\theta}\phi(\theta|\Omega_j)d\theta\end{aligned} \tag{A.5}$$

where $\frac{\partial k_j}{\partial \tilde{b}_j} = \frac{q_j}{\sqrt{1+2\phi\tilde{b}_j q_j}}$, and $\frac{\partial C}{\partial \tilde{b}_j} = \frac{1}{\tilde{b}_j} - \alpha\frac{1}{k_j}\frac{\partial k_j}{\partial \tilde{b}_j}$.

$$\begin{aligned}\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{q}_j} =& -[pdef_j]\,\psi c_d \frac{\partial k_j}{\partial q_j} - \left[\int_{-\infty}^{\infty} \phi(C|\theta, z_j)\frac{\partial C}{\partial q_j} d\epsilon_j \phi(\theta|\Omega_j)d\theta\right]\psi c_d k_j \\ & + \alpha k_j^{\alpha-1}\frac{\partial k_j}{\partial q_j}\int_{-\infty}^{\infty}\int_{C(\theta)}^{\infty}\Lambda(\theta)e^{\epsilon_j}\phi(\epsilon_j|\theta, z_j)d\epsilon_j e^{\theta}\phi(\theta|\Omega_j)d\theta\end{aligned} \tag{A.6}$$

where $\frac{\partial k_j}{\partial q_j} = \frac{\tilde{b}_j}{\sqrt{1+2\phi\tilde{b}_j q_j}}$, and $\frac{\partial C}{\partial q_j} = -\alpha\frac{1}{k_j}\frac{\partial k_j}{\partial q_j}$.

$$\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial \tilde{b}_j} = \left[(1-pdef_j) - (1-\psi pdef_j)\frac{q_j}{q^f}\right] + \tilde{b}_j\left[-\frac{\partial pdef_j}{\partial \tilde{b}_j} + \psi\frac{q_j}{q^f}\frac{\partial pdef_j}{\partial \tilde{b}_j}\right] \tag{A.7}$$

where

$$\frac{\partial pdef_j}{\partial \tilde{b}_j} = \left[\int_{-\infty}^{\infty} \phi(C|\theta, z_j)\frac{\partial C}{\partial \tilde{b}_j} d\epsilon_j \phi(\theta|\Omega_j)d\theta\right] \tag{A.8}$$

Finally,

$$\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial q_j} = +\tilde{b}_j\left[-\frac{\partial pdef_j}{\partial q_j} + \psi\frac{q_j}{q^f}\frac{\partial pdef_j}{\partial q_j} - (1-\psi pdef_j)\frac{1}{q^f}\right] \tag{A.9}$$

where

$$\frac{\partial pdef_j}{\partial q_j} = \left[\int_{-\infty}^{\infty} \phi(C|\theta, z_j)\frac{\partial C}{\partial q_j} d\epsilon_j \phi(\theta|\Omega_j)d\theta\right] \tag{A.10}$$

A.3 Proofs

Proposition 1. Assume no moral hazard and no investment adjustment cost $c_d = \psi = \phi = 0$. To simplify the exposition, I drop the subscript j. Use the definition of $q = \frac{1}{1+r}$ and $q\tilde{b} = k$. As a result, $C = \left(\frac{k^{1-\alpha}}{q\Lambda(\theta)}\right) - \theta$.

Foc 1 Consider the first first order condition in A.4.

$$q = q^f \left[1 - \int_{-\infty}^{\infty} \Phi_e(C(\theta)|z,\theta)\phi_\theta(\theta|\Omega)d\theta\right] \tag{A.11}$$

In steady state

$$q^* = q^f \left[1 - \Phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\right] \tag{A.12}$$

where x^* is the steady state value of variable x. Differentiating

$$dq = -\, q^f \Phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right) \int_{-\infty}^{\infty} \left[dC - dE[e|z,\theta]\right] \phi_\theta\left(\theta|z\right) d\theta \tag{A.13}$$

where $dC = (1-\alpha)]\hat{k} - \hat{q} - (\eta_{\Lambda(M),\theta} - 1)d\theta$, where $\eta_{\Lambda(M),\theta} \equiv -\frac{1}{\Lambda(M)}\Lambda'(M)M'(\theta)$, and $dE[\epsilon|z,\theta] = \frac{\partial E[\epsilon|\tilde{z}]}{\partial\theta}d\theta + \frac{\partial E[\epsilon|\tilde{z}]}{\partial z}dz$. Therefore

$$\begin{aligned} dq = -q^f \Phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right) \int_{-\infty}^{\infty} \left[(1-\alpha)]\hat{k} - \hat{q} - \eta_{\Lambda(M),\theta}d\theta - \frac{\partial E[\epsilon|\tilde{z}]}{\partial\theta}d\theta\right] \phi_\theta\left(\theta|z\right) d\theta \\ -\, q^f \Phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right) \int_{-\infty}^{\infty} \left[-\frac{\partial E[\epsilon|\tilde{z}]}{\partial z}dz = \theta\right] \phi_\theta\left(\theta|z\right) d\theta \end{aligned} \tag{A.14}$$

Denote $a \equiv ln(A)$ and notice that

$$\begin{aligned} E[a|z,\theta] &= \tilde{m}(z-\theta)+\theta \\ &= \frac{\partial E[\epsilon|\tilde{z}]}{\partial z} z + \frac{\partial E[\epsilon|\tilde{z}]}{\partial \theta}\theta + \theta \end{aligned}$$

Moreover, $\hat{M} \equiv \frac{dM}{M} = \frac{M'(\theta)d\theta}{M}$ and therefore

$$\begin{aligned} \eta_{\Lambda(M),\theta} d\theta &= -\frac{M}{\Lambda(M)}\Lambda'(M)\frac{M'(\theta)d\theta}{M} \\ &= \eta_{\Lambda,M}\hat{M} \end{aligned} \tag{A.15}$$

where $\eta_{\Lambda,M} \equiv \frac{\nu-\xi}{1-(1-\alpha)\xi}$. Substitute back and divided by steady state value

$$\hat{q} = \tilde{L}_1 \left\{ -(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M} E[\hat{M}|z] \right\} \tag{A.16}$$

where $\tilde{L}_1 = \frac{\phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)}{\left[1-\Phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)-\phi\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\right]}$.

Foc 2 Differentiate the second first order condition in A.4

$$\frac{d\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}}{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}} - \frac{d\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial \tilde{b}_j}}{\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial \tilde{b}_j}} = \frac{d\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial q_j}}{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial q_j}} - \frac{d\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial q_j}}{\frac{\partial E[w_{bank,j}|\Omega_j]}{\partial q_j}} \tag{A.17}$$

and let's see each term individually.

- From equation A.5, the derivative of expected firm manager payoff with respect to bond $\tilde{b}$ is given by

$$\begin{aligned} \frac{\partial E[w_{firm}|\Omega]}{\partial \tilde{b}} = &- \left[1 - \int_{-\infty}^{\infty} \Phi_e(C(\theta)|z,\theta)\phi_\theta(\theta|\Omega)d\theta\right] \\ &+ \alpha k_j^{\alpha-1} q \int_{-\infty}^{\infty} \Lambda(\theta) e^{\frac{Var[\epsilon|\theta,z]}{2}+E[\epsilon|\theta,z]} \times \\ &\times \Phi_\epsilon \left(\frac{Var[\epsilon|\theta,z]+E[\epsilon|\theta,z]-C(\theta)}{\sqrt{Var[\epsilon|\theta,z]}}\right) e^\theta \phi(\theta|\Omega_j)d\theta \end{aligned}$$

Differentiating,

$$\begin{aligned}
d\frac{\partial E[w_{firm}|\Omega]}{\partial \tilde{b}} =&\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\
&+\alpha k_j^{\alpha-1}q\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\Phi_\epsilon\left(\cdot\right)\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\
&-\alpha k_j^{\alpha-1}q\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\phi_\epsilon\left(\cdot\right)\left\{-\hat{q}+(1-\alpha)\,\hat{k}-E[a|z]-\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\
=&\left\{\alpha k_j^{\alpha-1}q\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\left[\Phi_\epsilon\left(\cdot\right)+\phi_\epsilon\left(\cdot\right)\right]+\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\right\}\times \\
&\times\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}
\end{aligned} \tag{A.18}$$

As a result,

$$\begin{aligned}
\frac{d\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}}{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}} =& \frac{\left\{\alpha k_j^{\alpha-1}q\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\left[\Phi_\epsilon\left(\cdot\right)+\phi_\epsilon\left(\cdot\right)\right]+\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\right\}}{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}}\times \\
&\times\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\
&=L_1\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}
\end{aligned} \tag{A.19}$$

where $L_1 \equiv \frac{\left\{\alpha k_j^{\alpha-1}q\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\left[\Phi_\epsilon(\cdot)+\phi_\epsilon(\cdot)\right]+\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\right\}}{\frac{\partial E[w_{firm,j}|\Omega_j]}{\partial \tilde{b}_j}}$.

- From equation A.6, the derivative of expected firm manager payoff with respect to bond price q is given by

$$\begin{aligned}
\frac{\partial E[d_{firm}|\Omega]}{\partial q} =&\alpha k_j^{\alpha-1}\frac{k}{q}\int_{-\infty}^{\infty}\Lambda(\theta)e^{\frac{Var[\epsilon|\theta,z]}{2}+E[\epsilon|\theta,z]}\times \\
&\times\Phi_\epsilon\left(\frac{Var[\epsilon|\theta,z]+E[\epsilon|\theta,z]-C(\theta)}{\sqrt{Var[\epsilon|\theta,z]}}\right)e^\theta\phi(\theta|\Omega_j)d\theta
\end{aligned} \tag{A.20}$$

Differentiating,

$$\begin{aligned} d\frac{\partial E[d_{firm}|\Omega]}{\partial q} =& \alpha k_j^{\alpha-1}\frac{k}{q}\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\left[\Phi_\epsilon\left(\cdot\right)+\phi_\epsilon\left(\cdot\right)\right]\times \\ & \times\left\{\hat{q}-\left(1-\alpha\right)\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\ & +\alpha k_j^{\alpha-1}\frac{k}{q}\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\Phi_\epsilon\left(\cdot\right)(\hat{k}-2\hat{q}) \end{aligned} \tag{A.21}$$

therefore

$$\begin{aligned} \frac{d\frac{\partial E[w_{firm,j}|\Omega]}{\partial q}}{\frac{\partial E[w_{firm,j}|\Omega]}{\partial q}} &= \frac{\alpha k_j^{\alpha-1}\frac{k}{q}\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}\left[\Phi_\epsilon\left(\cdot\right)+\phi_\epsilon\left(\cdot\right)\right]}{\frac{\partial E[w_{firm,j}|\Omega]}{\partial q}}\times \\ &\times\left\{\hat{q}-\left(1-\alpha\right)\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}+(\hat{k}-2\hat{q}) \\ &= L_2\left\{\hat{q}-\left(1-\alpha\right)\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}+(\hat{k}-2\hat{q}) \end{aligned} \tag{A.22}$$

where $L_2 \equiv \frac{\alpha k_j^{\alpha-1}\frac{k}{q}\Lambda e^{\frac{Var[\epsilon|\theta,z]}{2}}[\Phi_\epsilon(\cdot)+\phi_\epsilon(\cdot)]}{\frac{\partial E[w_{firm,j}|\Omega]}{\partial q}}$.

- From equation A.7, the derivative of the expected bank manager payoff with respect to bond $\tilde{b}_j$ is given by

$$\begin{aligned} \frac{\partial E[d_{bank}|\Omega^i]}{\partial\tilde{b}} =& \left[\left(1-\int_{-\infty}^{\infty}\Phi_e(C(\theta)|z,\theta)\phi_\theta(\theta|\Omega)d\theta\right)-\frac{q}{q^f}\right] \\ & -(1-\alpha)\int_{-\infty}^{\infty}\phi_e(C(\theta)|z,\theta)\phi_\theta(\theta|\Omega)d\theta \end{aligned}$$

Differentiating,

$$\begin{aligned} d\frac{\partial E[d_{bank}|\Omega]}{\partial\tilde{b}} =& -\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\left\{\hat{q}-\left(1-\alpha\right)\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \\ & -\frac{q}{q^f}\hat{q}-(1-\alpha)\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{C^*}{Var[\epsilon|\theta,z]}\times \\ & \times\left\{\hat{q}-\left(1-\alpha\right)\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} \end{aligned} \tag{A.23}$$

therefore

$$\frac{d\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}}{\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}} = \frac{\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)(1+(1-\alpha)\frac{C}{Var[\epsilon|\theta,z]})}{\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}}\times$$
$$\times\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} - \frac{\frac{q}{q^f}}{\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}}\hat{q} \tag{A.24}$$
$$= L_3\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\} - L_4\hat{q}$$

where $L_3 \equiv \frac{\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)(1+(1-\alpha)\frac{C}{Var[\epsilon|\theta,z]})}{\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}}\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}$ and $L_4 \equiv \frac{\frac{q}{q^f}}{\frac{\partial E[d_{bank}|\Omega]}{\partial \tilde{b}}}$.

- From equation A.9, the derivative of the expected bank manager payoff with respect to bond price q_j is given by

$$\frac{\partial E[d_{bank}|\Omega]}{\partial q} = \frac{k}{q}\left[\alpha\int_{-\infty}^{\infty}\phi_e(C(\theta)|z,\theta)\phi_\theta(\theta|\Omega)d\theta\frac{1}{q} - \frac{1}{q^f}\right]$$

differentiating,

$$d\frac{\partial E[d_{bank}|\Omega]}{\partial q} = \frac{k}{q}(\hat{k}-\hat{q})\left[\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{1}{q} - \frac{1}{q^f}\right] +$$
$$+\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{C^*}{Var[\epsilon|\theta,z]}\frac{1}{q}\times$$
$$\times\left\{\hat{q}-(1-\alpha)\,\hat{k}+E[a|z]+\eta_{\Lambda,M}E[\hat{M}|z]\right\}$$
$$-\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{1}{q}\hat{q}$$

Therefore

$$\frac{d\frac{\partial E[d_{bank}|\Omega]}{\partial q}}{\frac{\partial E[d_{bank}|\Omega]}{\partial q}} = (\hat{k} - \hat{q}) - \frac{\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)}{\frac{\partial E[d_{bank}|\Omega]}{\partial q}}\hat{q}$$
$$- \frac{\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{C^*}{Var[\epsilon|\theta,z]}}{\frac{\partial E[d_{bank}|\Omega]}{\partial q}}\left\{\hat{q} - (1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\}$$
$$= (\hat{k} - \hat{q}) - L_5\left\{\hat{q} - (1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\} - L_6\hat{q} \tag{A.25}$$

where $L_5 = \frac{\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)\frac{C^*}{Var[\epsilon|\theta,z]}}{\frac{\partial E[d_{bank}|\Omega]}{\partial q}}$ and $L_6 = \frac{\frac{k}{q}\alpha\phi_e\left(\frac{C^*}{\sqrt{Var[\epsilon|\theta,z]}}\right)}{\frac{\partial E[d_{bank}|\Omega]}{\partial q}}$.

Finally, substitute equations A.19, A.22, A.24, and A.25 in equation A.17 and get

$$\hat{q} = \frac{-(L_1 - L_2 - L_3 - L_5)\left\{-(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\}}{(L_1 - L_2 - L_3 - L_5 + L_4 + 1 - L_6)}$$
$$\hat{q} = \tilde{L}_2\left\{-(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\} \tag{A.26}$$

where $\tilde{L}_2 \equiv \frac{-(L_1 - L_2 - L_3 - L_5)}{(L_1 - L_2 - L_3 - L_5 + L_4 + 1 - L_6)}$.

Equilibrium Substitute equation A.16 in A.26

$$\tilde{L}_1\left\{-(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\} = \tilde{L}_2\left\{-(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\}$$
$$(\tilde{L}_1 - \tilde{L}_2)\left\{-(1-\alpha)\hat{k} + E[a|z] + \eta_{\Lambda,M}E[\hat{M}|z]\right\} = 0 \tag{A.27}$$

therefore, the stage-2 equilibrium k and q are given by

$$\hat{k} = \frac{1}{1-\alpha}(E[a|z] - \gamma E[\hat{M}|z])$$
$$\hat{q} = 0 \tag{A.28}$$

Where $\gamma \equiv -\eta_{\Lambda(M),M} = -\frac{\nu-\xi}{1-(1-\alpha)\xi}$. If $\nu < \xi$, then $\gamma > 0$. Therefore $\hat{r}_j \propto \hat{q} = 0$.

Since $M = \left\{ \left[\frac{w}{(1-\alpha)\xi\nu}\right]^{\frac{(1-\alpha)}{(1-\alpha)\xi-1}} \left[\int^N A_j k_j^\alpha dj\right]^{\frac{1}{\xi}} \right\}^{\frac{1-(1-\alpha)\xi}{1-(1-\alpha)\nu}}$, log deviation of M around the stochastic steady state equals

$$\hat{M} = \mu(\alpha\hat{K} + \theta)$$

where $\mu \equiv \frac{1}{\xi}\frac{1-(1-\alpha)\xi}{1-(1-\alpha)\nu} > 0$ and $\hat{K} = \int^j k_j dj$. One can write

$$\begin{aligned}
\hat{k} &= \frac{1}{1-\alpha}(E[a|z] - \gamma\mu E[\theta + \alpha\hat{K}|z]) \\
\hat{q} &= 0
\end{aligned}$$

Moreover, from A.11

$$\hat{q}_j = -\zeta\hat{p}(def_j|\Omega_j) = 0 \tag{A.29}$$

where $\zeta \equiv \frac{p^*(def|0)}{1-p^*(def|0)}$.

The expected level deviation of bank j profits from steady state conditioning on state θ equals

$$\begin{aligned}
E[w_{bank,j}|z_j,\theta] &= -p^*(def|0)\hat{p}(def_j|z_j,\theta) - \frac{q^*}{q_j}\hat{q}_j \\
&= -p^*(def|0)[\hat{p}(def_j|z_j,\theta) - E[\hat{p}(def_j|\Omega_j)|\theta]]
\end{aligned} \tag{A.30}$$

which is zero for each θ if $\theta \in \Omega_j$. ■

Proposition 2. Consider the global game when θ is observed

$$\hat{k} = \frac{1}{1-\alpha}E[a_j|z] - \frac{1}{1-\alpha}\gamma\mu\left(\theta + \alpha\hat{K}\right) \tag{A.31}$$

where $E[a_j|z_j,\theta] = \tilde{m}(z_j - \theta) + \theta$, where $z_j = a_j + \eta_j$ and $\tilde{m} = \frac{\sigma_e^2}{\sigma_e^2+\sigma_\eta^2}$. Aggregating across islands

$$
\begin{aligned}
K &= \frac{1}{1-\alpha}(1-\gamma\mu)\theta - \frac{\alpha}{1-\alpha}\gamma\mu K \\
K &= \frac{(1-\gamma\mu)}{1-\alpha+\alpha\gamma\mu}\theta
\end{aligned}
\tag{A.32}
$$

■

Proposition 3. Consider the global game when θ is not observed

$$
\hat{k} = \frac{1}{1-\alpha}E[a_j|z_j] - \frac{1}{1-\alpha}\gamma\mu\left(E[\hat{\theta}|z_j] + \alpha E[\hat{K}|z_j]\right) \tag{A.33}
$$

where $E[a_j|z_j] = mz_j$, where $m = \frac{\sigma_e^2+\sigma_\theta^2}{\sigma_e^2+\sigma_\theta^2+\sigma_\eta^2}$, and $E[\theta|z_j] = \delta z_j$ where $\delta = \frac{\sigma_e^2}{\sigma_e^2+\sigma_\theta^2+\sigma_\eta^2}$. Following Morris and Shin (2002), I guess the linear solution $k_j = \chi z_j$

$$
\begin{aligned}
k_j &= \frac{1}{1-\alpha}(m-\gamma\mu[1+\alpha\chi]\delta)z_j \\
\chi &= \frac{1}{1-\alpha}(m-\gamma\mu[1+\alpha\chi]\delta) \\
\chi &= \frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta} \\
K &= \frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta}\theta
\end{aligned}
\tag{A.34}
$$

■

Corollary 4. The loglinearized individual revenues $\hat{\pi}_j$ if $\theta \notin \Omega_j$ equals

$$
\begin{aligned}
\widehat{\pi}_j &= -\gamma\hat{M} + a_j + \alpha k_j \\
&= -\gamma\mu\left(\theta + \alpha\frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta}\theta\right) + a_j + \alpha k_j
\end{aligned}
\tag{A.35}
$$

Since $E[a_j|z_j] = mz_j$ and $E[\theta|z_j] = \delta z_j$,

$$\begin{aligned} E[\widehat{\pi}_j|z_j,\theta] - E[E[\widehat{\pi}_j|z_j]|\theta] =& E[a_j|z_j,\theta] - E[E[a_j|z_j]|\theta] - \gamma(\hat{M} - E[\hat{M}|z_j]) \\ =& \left[(1-m) - \gamma\mu(1-\delta)\left(1+\alpha\frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta}\right)\right]\theta \end{aligned} \tag{A.36}$$

It implies that average forecast errors are a positive function of θ if

$$\begin{aligned} (1-m) - \gamma\mu(1-\delta)\left(1+\alpha\frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta}\right) > 0 \\ (m-\gamma\mu\delta)(1-\alpha+\alpha\gamma\mu) > (1-\gamma\mu)(1-\alpha+\gamma\mu\alpha\delta) \end{aligned} \tag{A.37}$$

■

Corollary 5. Consider actual probability of default of firm j in dispersed information conditioning on aggregate shock θ: $p(def_j|z_j,\theta) \equiv \Phi_{e|\tilde{z}}(C(\theta))$. The first order approximation around the risky steady state is

$$\hat{p}(def_j|z_j,\theta) = \frac{\phi_{e|0}(C^*)}{\Phi_{e_0}(C^*)}\left[(1-\alpha)\hat{k}_j - \hat{q}_j + \gamma\hat{M} - E[a_j|z_j,\theta]\right] \tag{A.38}$$

Aggregating across islands

$$\begin{aligned} \hat{p}(def|z_j,\theta) &= \xi\left[(1-\alpha)\hat{K} - \hat{Q} + \gamma\hat{M} - \theta\right] \\ \hat{p}(def|z_j,\theta) &= \xi\left[(1-\alpha+\alpha\gamma\mu)\hat{K} - (1-\gamma\mu)\theta\right] \\ \hat{p}(def|z_j,\theta) &= \xi\left[(1-\alpha+\alpha\gamma\mu)\frac{(m-\gamma\mu\delta)}{1-\alpha+\gamma\mu\alpha\delta} - (1-\gamma\mu)\right]\theta \end{aligned} \tag{A.39}$$

Then it implies that $\frac{\partial\hat{p}(def|\theta)}{\partial\theta} > 0$ if

$$(m-\gamma\mu\delta)(1-\alpha+\alpha\gamma\mu) > (1-\gamma\mu)(1-\alpha+\gamma\mu\alpha\delta) \tag{A.40}$$

■

Corollary 6. Consider the logdeviation of perceived probability of default from steady state, meaning conditioning on info set $\Omega_j = \{z_j\}$.

$$\hat{p}(def_j|z_j) = \frac{\phi_{e|0}(C^*)}{\Phi_{e_0}(C^*)}\left[(1-\alpha)\,\hat{k}_j - \hat{q}_j + \gamma E[\hat{M}|z_j] - E[a_j|z_j]\right] \tag{A.41}$$

Consider the logdeviation of actual probability of default from steady state, meaning conditioning on info set $\Omega_j = \{z_j, \theta\}$.

$$\hat{p}(def_j|z_j,\theta) = \frac{\phi_{e|0}(C^*)}{\Phi_{e_0}(C^*)}\left[(1-\alpha)\,\hat{k}_j - \hat{q}_j + \gamma \hat{M} - E[a_j|z_j,\theta]\right] \tag{A.42}$$

The average bank profits equal the difference between the two

$$\begin{aligned} E[\tilde{\pi}_{bank,j}|z_j,\theta] &\propto -[\hat{p}(def_j|z_j,\theta) - E[\hat{p}(def_j|z_j)|\theta]] \\ &\propto -[E[a_j|z_j,\theta] - E[E[a_j|z_j]|\theta] - \gamma(M - E[M|z_j])] \end{aligned} \tag{A.43}$$

from the proof of corollary 4, it follow that average bank profits are a negative function of θ if

$$(m - \gamma\mu\delta)\,(1-\alpha+\alpha\gamma\mu) > (1-\gamma\mu)(1-\alpha+\gamma\mu\alpha\delta) \tag{A.44}$$

■

A.4 Equal bargaining power

In the baseline model I assume firm and bank managers decide loan quantity and prices in second stage and information in the first stage through Nash bargaining, with the firms retaining all bargaining power. This yields the standard implication that the price of the loan reflects only quantity of risk, with no changes in price of risk. I relax this assumption here by setting the same bargaining power on bank and firm.

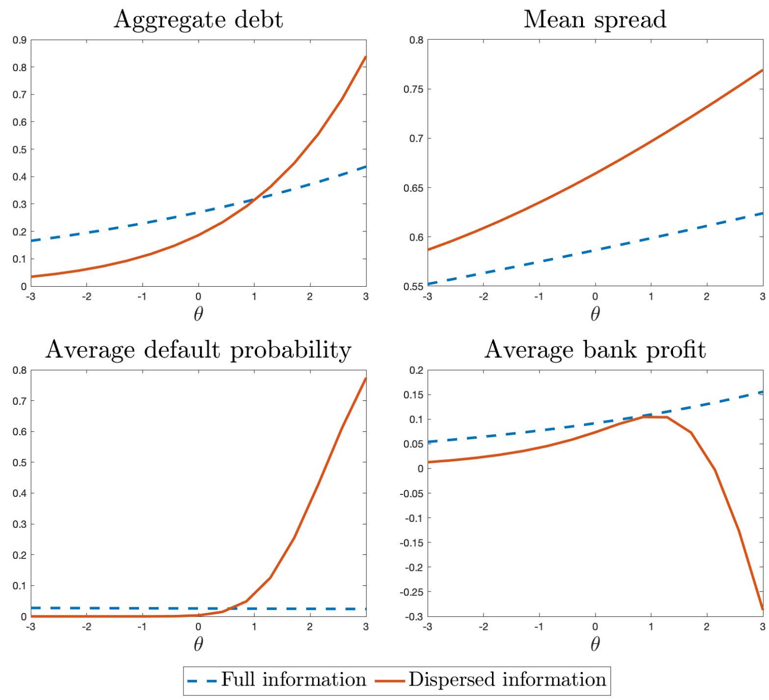


Figure A.1: Model with alternative calibration: full and dispersed information

Second stage The second-stage optimal k_j^* and q_j^* maximize

$$max_{q_j,b_j}\ (E[w_{firm}|\Omega_j])^{\beta}(E[w_{bank}|\Omega_j])^{1-\beta} \tag{A.45}$$

Figure A.1, and A.2 illustrate the equilibrium where $\beta = 0.5$. Differently from the baseline model, risk premium increases in booms even if risk declines, as the bank extract more profit from the firm. As a result, bank's profits increase in moderate booms, but decline for very large booms as the losses for mispricing of risk becomes larger than the rent extraction from the firm.

First stage Next , consider the same convex compensation structure 1.8 on the bank manager instead

$$w_{bank} = \begin{cases} b_j(1-q_j) - b^f(1-q^f) & \text{if } \Lambda(M)A_j{k_j}^{\alpha} \geq b_j \quad (repay) \\ (1-\alpha_b)[b_j(-q_j) - b^f(1-q^f)] & \text{if } \Lambda(M)A_j{k_j}^{\alpha} < b_j \quad (default) \end{cases} \tag{A.46}$$

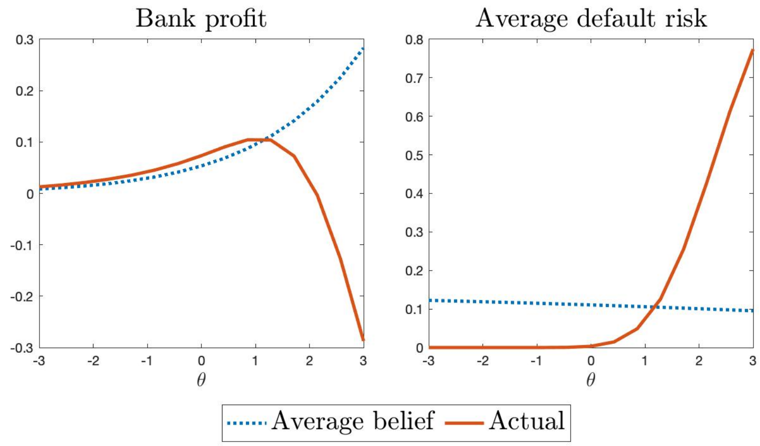


Figure A.2: Model with alternative calibration: actuals and beliefs

where α_b is the option holding of the bank manager. In the first stage the island decide to pay the information cost if

$$\begin{aligned} &(E[\pi^*_{firm}(\theta \in \Omega_j, \lambda)] - \beta c)^{\beta}(E[w^*_{bank}(\theta \in \Omega_j, \lambda)] - (1-\beta)c)^{1-\beta} \\ &\geq (E[\pi^*_{firm}(\theta \notin \Omega_j, \lambda)])^{\beta}(E[w^*_{bank}(\theta \notin \Omega_j, \lambda)])^{1-\beta} \end{aligned} \tag{A.47}$$

where I assume that bank and firm split the information cost c according to their bargaining power β as well. Figure A.3 reports the equilibrium information λ for different values of bank manager compensation convexity (assuming no convexity on firm manager's compensation). Higher moral hazard incentives on bank manager aalso reduces optimal information choice.

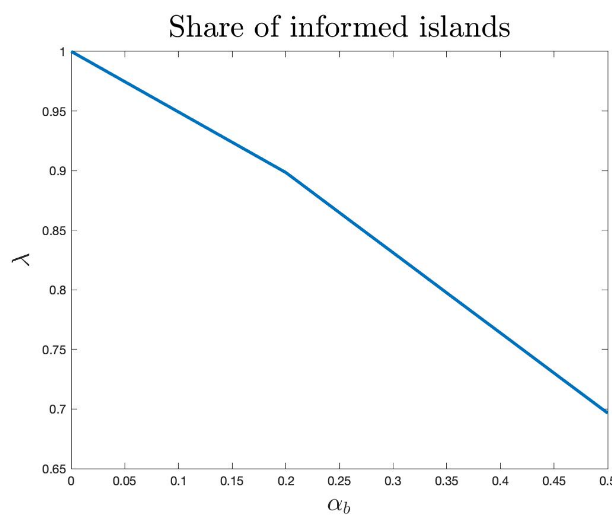


Figure A.3: Moral hazard and information with alternative calibration

Appendix B

Biased Surveys

B.1 Variable definitions

All variables come from the Survey of Professional Forecasters, collected by the Federal Reserve Bank of Philadelphia. All surveys are collected around the 3rd week of the middle month in the quarter. In this section, x_t indicate the actual value and $F_t x_t + h$ the forecast provided in t about horizon h. All actual values of macroeconomic series (1-12) use the first release level, which are available to forecasters in the following quarter. We transform the macroeconomic level in growth. The series are constructed similarly as Bordalo et al. (2020)

1. NGDP

   - Variable: nominal GDP.
   - Question: The level of nominal GDP in the current quarter and the next 4 quarters.
   - Forecast: Nominal GDP growth from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
   - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$

- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

2. RGDP

- Variable: real GDP.
- Question: The level of real GDP in the current quarter and the next 4 quarters.
- Forecast: real GDP growth from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

3. PGDP

- Variable: GDP deflator.
- Question: The level of GDP deflator in the current quarter and the next 4 quarters.
- Forecast: GDP price deflator inflation from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

4. CPI

- Variable: Consumer Price Index.
- Question: CPI growth rate in the current quarter and the next 4 quarters.
- Forecast: CPI inflation from end of quarter $t-1$ to end of quarter $t+3$: $F_t(z_t/4+1) * F_t(z_{t+1}/4+1) * F_t(z_{t+2}/4+1) * F_t(z_{t+3}/4+1)$, where z is the annualized quarterly CPI inflation in quarter t.

- Revision: $F_t(z_t/4+1) * F_t(z_{t+1}/4+1) * F_t(z_{t+2}/4+1) * F_t(z_{t+3}/4+1) - F_{t-1}(z_t/4+1) * F_{t-1}(z_{t+1}/4+1) * F_{t-1}(z_{t+2}/4+1) * F_{t-1}(z_{t+3}/4+1)$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$. Real time data is not available before 1994Q3. For actual periods prior to this date, we use data published in 1994Q3 to measure the actual outcome.

5. RCONSUM

- Variable: Real consumption.
- Question: The level of real consumption in the current quarter and the next 4 quarters.
- Forecast: GDP price deflator inflation from end of quarter $t-1$ to end of quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

6. INDPROD

- Variable: Industrial production index.
- Question: The average level of the industrial production index in the current quarter and the next 4 quarters.
- Forecast: Growth of the industrial production index from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

7. RNRESIN

- Variable: Real non-residential investment.

- Question: The level of real non-residential investment in the current quarter and the next 4 quarters.
- Forecast: Growth of real non-residential investment from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

8. RRESIN

   - Variable: Real residential investment.
   - Question: The level of real residential investment in the current quarter and the next 4 quarters.
   - Forecast: Growth of real residential investment from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
   - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
   - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

9. RGF

   - Variable: Real federal government consumption.
   - Question: The level of real federal government consumption in the current quarter and the next 4 quarters.
   - Forecast: Growth of real federal government consumption from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
   - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
   - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

10. RGSL

- Variable: Real state and local government consumption.
- Question: The level of real state and local government consumption in the current quarter and the next 4 quarters.
- Forecast: Growth of real state and local government consumption from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
- Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
- Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$

11. HOUSING
    - Variable: Housing starts.
    - Question: The level of housing starts in the current quarter and the next 4 quarters.
    - Forecast: Growth of housing starts from quarter $t-1$ to quarter $t+3$: $\frac{F_t x_{t+3}}{x_{t-1}} - 1$
    - Revision: $\frac{F_t x_{t+3}}{x_{t-1}} - \frac{F_{t-1} x_{t+3}}{F_{t-1} x_{t-1}}$
    - Actual: $\frac{x_{t+3}}{x_{t-1}} - 1$
12. UNEMP
    - Variable: Unemployment rate.
    - Question: The level of average unemployment rate in the current quarter and the next 4 quarters.
    - Forecast: Average quarterly unemployment rate in quarter $t+3$: $F_t x_{t+3}$
    - Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
    - Actual: x_{t+3}
13. TB3M

- Variable: 3-month Treasury rate.
- Question: The level of average 3-month Treasury rate in the current quarter and next 4 quarters.
- Forecast: Average quarterly 3-month Treasury rate in quarter $t+3$: $F_t x_{t+3}$
- Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
- Actual: x_{t+3}

14. TN10Y

- Variable: 10-year Treasury rate.
- Question: The level of average 10-year Treasury rate in the current quarter and next 4 quarters.
- Forecast: Average quarterly 10-year Treasury rate in quarter $t+3$: $F_t x_{t+3}$
- Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
- Actual: x_{t+3}

15. AAA

- Variable: AAA corporate bond rate.
- Question: The level of average AAA corporate bond rate in the current quarter and next 4 quarters.
- Forecast: Average quarterly AAA corporate bond rate in quarter $t+3$: $F_t x_{t+3}$
- Revision: $F_t x_{t+3} - F_{t-1} x_{t+3}$
- Actual: x_{t+3}

B.2 Proofs

Proposition 4. Let $\hat{x}_{t,t} \equiv \tilde{E}_t[x_t]$. From 2.3

$$\underbrace{(x_t - \bar{\hat{x}}_{t,t})}_{\bar{fe}_{t,t}} = \frac{1-G}{G}(\underbrace{\hat{\hat{x}}_{t,t} - \rho\hat{\hat{x}}_{t-1,t-1}}_{\bar{fr}_{t,t}}) - \frac{G_1}{G}e_t \tag{B.1}$$

therefore by running CG regression 2.4, the regressor $\bar{fr}_{t,t}$ is correlated with the unobservable error. The resulting $\hat{\beta}_{CG}$ is equal to:

$$\begin{aligned}
\beta_{CG} &= \frac{1-G}{G} + \frac{cov(G[x_t - \rho\bar{\hat{x}}_{t-1,t-1}] + G_1 e_t, -\frac{G_1}{G}e_t)}{var(G[x_t - \rho\bar{\hat{x}}_{t-1,t-1}] + G_1 e_t)} \\
&= \frac{1-G}{G} - \frac{\frac{G_1^2}{C}\nu^{-1}}{G^2 var(x_t - \rho\bar{\hat{x}}_{t-1,t-1}) + G_1^2\nu^{-1}} \\
&= \frac{var(x_t - \rho\bar{\hat{x}}_{t-1,t-1}) - [Gvar(x_t - \rho\bar{\hat{x}}_{t-1,t-1}) + \frac{G_1^2}{G}\nu^{-1}]}{Gvar(x_t - \rho\bar{\hat{x}}_{t-1,t-1}) + \frac{G_1^2}{G}\nu^{-1}}
\end{aligned} \tag{B.2}$$

but $var(x_t - \rho\bar{\hat{x}}_{t-1,t-1}) \neq \bar{\Sigma} \equiv var(x_t - \rho\bar{x}_{t-1,t-1})$. In steady state

$$\begin{aligned}
x_{t+1} - \rho\bar{\hat{x}}_{t,t} &= \rho(x_t - \bar{\hat{x}}_{t,t}) + u_{t+1} \\
\hat{\Sigma} &= \rho^2\bar{\hat{\Phi}} + \xi^{-1}
\end{aligned} \tag{B.3}$$

where $\bar{\hat{\Phi}} = var(x_t - \hat{x}_{t,t})$.

$$\begin{aligned}
x_t - \bar{\hat{x}}_{t,t} &= (1-G)[x_t - \rho\bar{\hat{x}}_{t-1,t-1}] - G_1 e_t \\
\bar{\hat{\Phi}} &= (1-G)^2\bar{\hat{\Sigma}} + G_1^2\nu^{-1}
\end{aligned} \tag{B.4}$$

Substitute and solve for $\hat{\Sigma}$

$$\bar{\hat{\Sigma}} = \frac{\rho^2[C_1^2\nu^{-1}] + \xi^{-1}}{1 - \rho^2(1-G)^2} \tag{B.5}$$

■

Corollary 7. With rational expectation, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with $\Sigma \equiv var(x_t - E_t^i[x_t])$ and $\bar{\Sigma} \equiv var(x_t - \bar{E}_t[x_t])$.

- From B.2, it follows that if $\nu = 0$, $G_1 = 0$ and $\beta_{CG} = \frac{1-G_2}{G_2}$. Moreover, if $\tau \to \infty$, $G_2 = 1$ and $\beta_{CG} = 0$.
- From B.2, it follows that If $\tau = 0$, $\Sigma = \bar{\Sigma}$, $G = G_1$ and $\frac{1-G}{G} = \frac{\Sigma^{-1}}{\nu^{-1}}$. Therefore $\beta_{CG} = 0$.

■

Proposition 5. From B.2, $\frac{1}{1+\beta_{CG}}$ is given by

$$\begin{aligned} \frac{1}{1+\beta_{CG}} &= \frac{1}{1 + \frac{1-G}{G} - \frac{\frac{G_1^2}{G}\nu^{-1}}{G^2\hat{\bar{\Sigma}}+G_1^2\nu^{-1}}} \\ &= G\left(\frac{G^2\hat{\bar{\Sigma}} + G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}}\right) > 0 \end{aligned} \tag{B.6}$$

which is equal to G if $G_1^2 = 0$. Subtracting the actual gain G

$$\begin{aligned} G\left(\frac{G^2\hat{\bar{\Sigma}} + G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}} - 1\right) \\ G\left(\frac{G_1^2\nu^{-1}}{G^2\hat{\bar{\Sigma}}}\right) > 0 \end{aligned} \tag{B.7}$$

■

Proposition 6. Let $\hat{x}_{t,t} \equiv \tilde{E}_t[x_t]$ and $x_{t,t} \equiv E_t[x_t]$. From 2.3

$$\begin{aligned} \hat{x}^i_{t,t} &= \rho\hat{x}^i_{t-1,t-1} + G(x_t - \rho\hat{x}^i_{t-1,t-1}) + G_1 e_t + G_2\eta^i_t \\ \hat{x}^i_{t,t} &= \rho\hat{x}^i_{t-1,t-1} + Gx_t - G\hat{x}^i_{t,t} + G(\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1}) + G_1 e_t + G_2\eta^i_t \\ (1-G)(\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1}) &= +G(x_t - \hat{x}^i_{t,t}) + G_1 e_t + G_2\eta^i_t \\ \underbrace{(x_t - \hat{x}^i_{t,t})}_{fe^i_{t,t}} &= \frac{1-G}{G}\underbrace{(\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1})}_{fr^i_{t,t}} - \frac{G_1}{G}e_t - \frac{G_2}{G}\eta^i_t \end{aligned} \tag{B.8}$$

therefore by running BGMS regression 2.11, the regressor $fr^i_{t,t}$ is correlated with the unobservable error. The resulting $\hat{\beta}_{BGMS}$ is equal to:

$$\begin{aligned}
\beta_{BGMS} &= \frac{1-G}{G} + \frac{cov(G[x_t - \rho\hat{x}^i_{t-1,t-1}] + G_1 e_t + G_2\eta^i_t, -\frac{G_1}{G}e_t - \frac{G_2}{G}\eta^i_t)}{var(G[x_t - \rho\hat{x}^i_{t-1,t-1}] + G_1 e_t + G_2\eta^i_t)} \\
&= \frac{1-G}{G} - \frac{\frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}{G^2 var(x_t - \rho\hat{x}^i_{t-1,t-1}) + G_1^2\nu^{-1} + G_2^2\tau^{-1}} \\
&= \frac{var(x_t - \rho\hat{x}^i_{t-1,t-1}) - [G var(x_t - \rho\hat{x}^i_{t-1,t-1}) + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}]}{G var(x_t - \rho\hat{x}^i_{t-1,t-1}) + \frac{G_1^2}{G}\nu^{-1} + \frac{G_2^2}{G}\tau^{-1}}
\end{aligned} \tag{B.9}$$

but $var(x_t - \rho\hat{x}^i_{t-1,t-1}) \neq \Sigma \equiv var(x_t - \rho x^i_{t-1,t-1})$. In steady state

$$\begin{aligned}
x_{t+1} - \rho\hat{x}^i_{t,t} =& \rho(x_t - \hat{x}^i_{t,t}) + u_{t+1} \\
\hat{\Sigma} =& \rho^2\hat{\Phi} + \xi^{-1}
\end{aligned} \tag{B.10}$$

where $\hat{\Phi} = var(x_t - \hat{x}^i_{t,t})$.

$$\begin{aligned}
x_t - \hat{x}^i_{t,t} =& (1-G)[x_t - \rho\hat{x}^i_{t-1,t-1}] - G_1 e_t - G_2\eta^i_t \\
\hat{\Phi} =& (1-G)^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}
\end{aligned} \tag{B.11}$$

Substitute and solve for $\hat{\Sigma}$

$$\hat{\Sigma} = \frac{\rho^2[G_1^2\nu^{-1} + G_2^2\tau^{-1}] + \xi^{-1}}{1 - \rho^2(1-G)^2} \tag{B.12}$$

■

Corollary 8. With rational expectation, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with

$\Sigma \equiv var(x_t - E_t^i[x_t])$. Define $\chi = \Sigma^{-1}$. From B.9

$$\begin{aligned}
\beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{\frac{1}{\chi+\nu+\tau}}{\frac{1}{(\chi+\nu+\tau)^2}[(\nu+\tau)^2\chi^{-1}+\nu+\tau]} \\
\beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{(\chi+\nu+\tau)\chi}{(\nu+\tau)^2+\nu\chi+\tau\chi} \\
\beta_{BGMS} &= \frac{\chi}{\nu+\tau} - \frac{(\chi+\nu+\tau)\chi}{(BGMS+\nu+\tau)(\nu+\tau)} \\
\beta_{BGMS} &= 0
\end{aligned} \tag{B.13}$$

■

Proposition 7. Let $\hat{x}_{t,t} \equiv \tilde{E}_t[x_t]$. From 2.3

$$\begin{aligned}
\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1} &= G(x_t - \rho\hat{x}^i_{t-1,t-1}) + G_1 e_t + G_2\eta^i_t \\
&= G_2(x_t - \rho\hat{x}^i_{t-1,t-1}) + G_1(g_t - \rho\hat{x}^i_{t-1,t-1}) + G_2\eta^i_t \\
&= G_2(x_t - \hat{x}^i_{t,t}) + G_2(\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1}) + G_1(g_t - \rho\hat{x}^i_{t-1,t-1}) + G_2\eta^i_t \\
\underbrace{(x_t - \hat{x}^i_{t,t})}_{fe^i_{t,t}} &= \frac{1-G_2}{G_2}\underbrace{(\hat{x}^i_{t,t} - \rho\hat{x}^i_{t-1,t-1})}_{fr^i_{t,t}} - \frac{G_1}{G_2}\underbrace{(g_t - \rho\hat{x}^i_{t-1,t-1})}_{pi^i_{t,t}} - \eta^i_t
\end{aligned} \tag{B.14}$$

Write regression 2.13 as

$$fe^i_{t,t} = X\beta + err^i_t \tag{B.15}$$

where $X = [fr^i_{t,t} \quad pi^i_{t,t}]$ and $\beta = \begin{bmatrix}\beta_1 \\ \beta_2\end{bmatrix}$.

$$\hat{\beta} = \beta + \Sigma^{-1}_{XX}\Sigma_{Xu} \tag{B.16}$$

where

$$
\begin{aligned}
\Sigma_{XX} &= \begin{bmatrix} var(fr^i_{t,t}) & cov(fr^i_{t,t}, pi^i_{t,t}) \\ cov(fr^i_{t,t}, pi^i_{t,t}) & var(pi^i_{t,t}) \end{bmatrix} \\
\Sigma_{XX}^{-1} &= \frac{1}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \begin{bmatrix} var(pi^i_{t,t}) & -cov(fr^i_{t,t}, pi^i_{t,t}) \\ -cov(fr^i_{t,t}, pi^i_{t,t}) & var(fr^i_{t,t}) \end{bmatrix} \\
\Sigma_{Xu} &= \begin{bmatrix} cov(fr^i_{t,t}, err^i) \\ cov(pi^i_{t,t}, err^i) \end{bmatrix} \\
\hat{\beta} &= \beta + \Sigma_{XX}^{-1}\Sigma_{Xu} = \beta + \begin{bmatrix} \frac{var(pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\ \frac{-cov(fr^i_{t,t}, pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \end{bmatrix}
\end{aligned}
\tag{B.17}
$$

and

$$
\begin{aligned}
var(fr^i_{t,t}) &= [G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}] \\
var(pi^i_{t,t}) &= \hat{\Sigma} + \nu^{-1} \\
cov(fr^i_{t,t}, pi^i_{t,t}) &= [G\hat{\Sigma} + G_1\nu^{-1}] \\
cov(fr^i_{t,t}, err^i) &= -G_2\tau^{-1} \\
cov(pi^i_{t,t}, err^i) &= 0
\end{aligned}
\tag{B.18}
$$

where $\hat{\chi} = \hat{\Sigma}^{-1}$ therefore

$$
\begin{aligned}
\hat{\beta}_1 &= \frac{1 - G_2}{G_2} + \frac{var(pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\
&= \frac{1 - G_2}{G_2} - \frac{(\hat{\Sigma} + \nu^{-1})G_2\frac{1}{\tau}}{(\hat{\Sigma} + \nu^{-1})(G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\hat{\Sigma} + G_1\nu^{-1})^2}
\end{aligned}
\tag{B.19}
$$

and

$$\begin{aligned}
\hat{\beta}_2 &= -\frac{G_1}{G_2} + \frac{-cov(fr^i_{t,t}, pi^i_{t,t})cov(fr^i_{t,t}, err)}{var(fr^i_{t,t})var(pi^i_{t,t}) - cov(fr^i_{t,t}, pi^i_{t,t})^2} \\
&= -\frac{G_1}{G_2} + \frac{(G\hat{\Sigma} + G_1\nu^{-1})G_2\frac{1}{\tau}}{(\hat{\Sigma} + \nu^{-1})(G^2\hat{\Sigma} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\hat{\Sigma} + G_1\nu^{-1})^2}
\end{aligned} \tag{B.20}$$

■

Corollary 9. With rational expectation, $G_1 = \frac{\nu}{\tau+\nu+\Sigma^{-1}}$ and $G_2 = \frac{\tau}{\tau+\nu+\Sigma^{-1}}$, with $\Sigma \equiv var(x_t - E^i_t[x_t])$. Define $\chi = \Sigma^{-1}$.

From B.19

$$\begin{aligned}
\beta_1 &= \frac{\chi+\nu}{\tau} - \frac{(\chi^{-1} + \nu^{-1})G_2\frac{1}{\tau}}{(\chi^{-1} + \nu^{-1})(G^2\chi^{-1} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\chi^{-1} + G_1\nu^{-1})^2} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\frac{1}{\chi+\nu+\tau})(\frac{\nu+\chi}{\nu\chi})}{(\frac{1}{\chi+\nu+\tau})^2[(\frac{\nu+\chi}{\nu\chi})(\frac{(\nu+\tau)^2}{\chi} + \nu + \tau) - (\frac{(\nu+\tau)}{\chi}) + 1)^2]} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\frac{\nu+\chi}{\nu\chi})(\chi+\nu+\tau)}{(\frac{\nu+\chi}{\nu\chi})(\nu+\tau)(\chi+\nu+\tau)\frac{1}{\chi} - (\chi+\nu+\tau)^2\frac{1}{\chi^2}} \\
&= \frac{\chi+\nu}{\tau} - \frac{(\nu+\chi)}{(\chi+\nu+\tau) - (\nu+\chi)} \\
&= 0
\end{aligned} \tag{B.21}$$

While from B.20

$$\begin{aligned}
\hat{\beta}_2 &= -\frac{\nu}{\tau} + \frac{(G\chi^{-1} + G_1\nu^{-1})G_2\frac{1}{\tau}}{(\chi^{-1} + \nu^{-1})(G^2\chi^{-1} + G_1^2\nu^{-1} + G_2^2\tau^{-1}) - (G\chi^{-1} + G_1\nu^{-1})^2} \\
&= -\frac{\nu}{\tau} + \frac{(\frac{1}{\chi+\nu+\tau})^2(\frac{\nu+\tau}{\chi} + 1)}{(\frac{1}{\chi+\nu+\tau})^2[(\frac{\nu+\chi}{\nu\chi})(\frac{(\nu+\tau)^2}{\chi} + \nu + \tau) - (\frac{(\nu+\tau)}{\chi}) + 1)^2]} \\
&= -\frac{\nu}{\tau} + \frac{(\chi+\nu+\tau)\frac{1}{\chi}}{(\frac{\nu+\chi}{\nu\chi})(\nu+\tau)(\chi+\nu+\tau)\frac{1}{\chi} - (\chi+\nu+\tau)^2\frac{1}{\chi^2}} \\
&= -\frac{\nu}{\tau} + \frac{\nu\chi}{(\nu+\chi)(\nu+\tau) - (\nu+\tau+\chi)\nu} \\
&= 0
\end{aligned} \tag{B.22}$$

■

Proposition 8. From 2.19, using $\delta = \delta_1 + \delta_2$

$$
\begin{aligned}
\hat{x}^i &= \mu + \delta(x - \mu) + \delta_1 e + \delta_2 \eta^i \\
\hat{x}^i &= \mu + \delta x - \delta \hat{x}^i + \delta(\hat{x}^i - \mu) + \delta_1 e + \delta_2 \eta^i \\
(1-\delta)(\hat{x}^i - \mu) &= +\delta(x - \hat{x}^i) + \delta_1 e + \delta_2 \eta^i \\
(\underbrace{x - \hat{x}^i}_{fe^i}) &= \frac{1-\delta}{\delta}(\underbrace{\hat{x}^i - \mu}_{fr^i}) - \frac{\delta_1}{\delta} e - \frac{\delta_2}{\delta} \eta^i
\end{aligned}
\tag{B.23}
$$

therefore by running BGMS regression 2.11, the regressor $fr^i = \hat{x}^i - \mu$ is correlated with the unobservable error. The resulting $\hat{\beta}_{BGMS}$ is equal to:

$$
\begin{aligned}
\hat{\beta}_{BGMS} &= \frac{1-\delta}{\delta} + \frac{cov(\hat{x}^i - \mu, -\frac{\delta_1}{\delta} e - \frac{\delta_2}{\delta}\eta^i)}{var(\hat{x}^i - \mu)} \\
&= \frac{1-\delta}{\delta} + \frac{cov(\delta(x_t - \mu) + \delta_1 e + \delta_2 \eta^i, -\frac{\delta_1}{\delta} e - \frac{\delta_2}{\delta}\eta^i)}{var(\delta(x_t - \mu) + \delta_1 e + \delta_2 \eta^i)} \\
&= \frac{1-\delta}{\delta} + \frac{-\frac{\delta_1^2}{\delta}\nu^{-1} - \frac{\delta_2^2}{\delta}\tau^{-1}}{\delta^2 \chi^{-1} + \delta_1^2 \nu^{-1} + \delta_2^2 \tau^{-1}}
\end{aligned}
\tag{B.24}
$$

substitute for δ_1 and δ_2

$$
\begin{aligned}
\hat{\beta}_{BGMS} = & \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)\gamma_1 + \gamma_2} \\
& - \frac{\frac{(1-\lambda)^2}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_1^2}{(1-\lambda)\gamma_1+\gamma_2}\nu^{-1} + \frac{1}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_2^2}{(1-\lambda)\gamma_1+\gamma_2}\tau^{-1}}{\frac{1}{[(1-\lambda)+\lambda\gamma_2]^2}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1} + (1-\lambda)^2\gamma_1^2\nu^{-1} + \gamma_2^2\tau^{-1})}
\end{aligned}
\tag{B.25}
$$

use definition of γ_1 and γ_2

$$\begin{aligned}
\hat{\beta}_{BGMS} &= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{\frac{1}{(1-\lambda)\nu+\tau}[(1-\lambda)^2\nu+\tau]}{\frac{1}{(1-\lambda)(\nu+\chi)+\tau}([[(1-\lambda)\nu+\tau]^2\chi^{-1}+(1-\lambda)^2\nu+\tau)} \\
&= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{[(1-\lambda)(\nu+\chi)+\tau][(1-\lambda)^2\nu+\tau]\chi}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)} \\
&= \frac{\chi\{(1-\lambda)[(1-\lambda)\nu+\tau]^2-[(1-\lambda)\nu+\tau][(1-\lambda)^2\nu+\tau]\}}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)} \\
&= \frac{-\lambda\tau\chi[(1-\lambda)\nu+\tau]}{[(1-\lambda)\nu+\tau]([[(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)} \\
&\frac{-\lambda\tau\chi}{([[(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)} < 0
\end{aligned} \tag{B.26}$$

which is negative as long as $0 < \lambda < 1$. ■

Proposition 9. From 2.19

$$\begin{aligned}
\bar{\hat{x}} &= \mu + \delta(x-\mu) + \delta_1 e \\
\bar{\hat{x}} &= \mu + \delta x - \delta\bar{\hat{x}} + \delta(\bar{\hat{x}}-\mu) + \delta_1 e \\
(1-\delta)(\bar{\hat{x}}-\mu) &= +\delta(x-\bar{\hat{x}}) + \delta_1 e \\
(\underbrace{x-\bar{\hat{x}}}_{fe^i}) &= \frac{1-\delta}{\delta}(\underbrace{\bar{\hat{x}}-\mu}_{fr^i}) - \frac{\delta_1}{\delta}e
\end{aligned} \tag{B.27}$$

therefore by running CG regression 2.4, the regressor $\bar{fr} = \bar{\hat{x}} - \mu$ is correlated with the unobservable error. The resulting $\hat{\beta}_{CG}$ is equal to:

$$\begin{aligned}
\hat{\beta}_{CG} &= \frac{1-\delta}{\delta} + \frac{cov(\bar{\hat{x}}-\mu, -\frac{\delta_1}{\delta}e)}{var(\bar{\hat{x}}-\mu)} \\
&= \frac{1-\delta}{\delta} + \frac{cov(\delta(x_t-\mu)+\delta_1 e, -\frac{\delta_1}{\delta}e)}{var(\delta(x_t-\mu)+\delta_1 e)} \\
&= \frac{1-\delta}{\delta} + \frac{-\frac{\delta_1^2}{\delta}\nu^{-1}}{\delta^2\chi^{-1}+\delta_1^2\nu^{-1}}
\end{aligned} \tag{B.28}$$

substitute for δ_1 and δ_2

$$\hat{\beta}_{CG} = \frac{(1-\lambda)(1-\gamma_1-\gamma_2)}{(1-\lambda)\gamma_1+\gamma_2} - \frac{\frac{(1-\lambda)^2}{(1-\lambda)+\lambda\gamma_2}\frac{\gamma_1^2}{(1-\lambda)\gamma_1+\gamma_2}\nu^{-1}}{\frac{1}{[(1-\lambda)+\lambda\gamma_2]^2}([(1-\lambda)\gamma_1+\gamma_2]^2\chi^{-1}+(1-\lambda)^2\gamma_1^2\nu^{-1})} \tag{B.29}$$

use definition of γ_1 and γ_2

$$\begin{aligned}
\hat{\beta}_{CG} &= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{\frac{1}{(1-\lambda)\nu+\tau}[(1-\lambda)^2\nu]}{\frac{1}{(1-\lambda)(\nu+\chi)+\tau}([(1-\lambda)\nu+\tau]^2\chi^{-1}+(1-\lambda)^2\nu)} \\
&= \frac{(1-\lambda)\chi}{(1-\lambda)\nu+\tau} - \frac{[(1-\lambda)(\nu+\chi)+\tau](1-\lambda)^2\nu\chi}{[(1-\lambda)\nu+\tau]([(1-\lambda)\nu+\tau]^2+(1-\lambda)^2\nu\chi)} \\
&= \frac{(1-\lambda)\tau\chi[(1-\lambda)\nu+\tau]}{[(1-\lambda)\nu+\tau]([(1-\lambda)\nu+\tau]^2+(1-\lambda)^2\nu\chi)} \\
&\frac{(1-\lambda)\tau\chi}{([(1-\lambda)\nu+\tau]^2+(1-\lambda)^2\nu\chi)} > 0
\end{aligned} \tag{B.30}$$

which is positive as long as $0 < \lambda < 1$. If $\lambda = 1$, it is zero. ■

Proposition 10. From 2.19

$$\begin{aligned}
x^i &= \mu + \delta_1(y-\mu) + \delta_2(x-\mu) + \delta_2\eta^i \\
x^i &= \mu + \delta_1(y-\mu) + \delta_2 x_t - \delta_2 x^i + \delta_2(\hat{x}^i-\mu) + \delta_2\eta^i \\
(1-\delta_2)(\hat{x}^i-\mu) &= \delta_1(y-\mu) + \delta_2(x-\hat{x}^i) + \delta_2\eta^i \\
(\underbrace{x-\hat{x}^i}_{fe^i}) &= \frac{1-\delta_2}{\delta_2}(\underbrace{\hat{x}^i-\mu}_{fr^i}) - \frac{\delta_1}{\delta_2}(\underbrace{g-\mu}_{pi}) - \eta^i
\end{aligned} \tag{B.31}$$

write regression 2.13 as

$$fe^i = X\beta + err^i \tag{B.32}$$

where $X = [fr^i \quad pi]$ and $\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$.

$$\hat{\beta} = \beta + \Sigma_{XX}^{-1}\Sigma_{Xu} \tag{B.33}$$

where

$$\begin{aligned}
\Sigma_{XX} &= \begin{bmatrix} var(fr^i) & cov(fr^i, pi) \\ cov(fr^i, pi) & var(pi) \end{bmatrix} \\
\Sigma_{XX}^{-1} &= \frac{1}{var(fr^i)var(pi) - cov(fr^i, pi)^2} \begin{bmatrix} var(pi) & -cov(fr^i, pi) \\ -cov(fr^i, pi) & var(fr^i) \end{bmatrix} \\
\Sigma_{Xu} &= \begin{bmatrix} cov(fr^i, err^i) \\ cov(pi, err^i) \end{bmatrix} \\
\hat{\beta} &= \beta + \Sigma_{XX}^{-1}\Sigma_{Xu} = \beta + \begin{bmatrix} \frac{var(pi)cov(fr,err)}{var(fr)var(pi) - cov(fr,pi)^2} \\ \frac{-cov(fr,pi)cov(fr,err)}{var(fr)var(pi) - cov(fr,pi)^2} \end{bmatrix}
\end{aligned} \tag{B.34}$$

and

$$\begin{aligned}
var(fr^i) &= \delta^2\chi^{-1} + \delta_1^2\nu^{-1} + \delta_2^2\tau^{-1} \\
var(pi) &= \chi^{-1} + \nu^{-1} \\
cov(fr^i, pi) &= \delta\chi^{-1} + \delta_1\nu^{-1} \\
cov(fr^i, err^i) &= -\delta_2\tau^{-1} \\
cov(pi, err^i) &= 0
\end{aligned} \tag{B.35}$$

$$\begin{aligned}
\hat{\beta}_1 &= \frac{1-\delta_2}{\delta_2} + \frac{var(pi)cov(fr,err)}{var(fr)var(pi)-cov(fr,pi)^2} \\
&= (1-\lambda)\frac{1-\gamma_2}{\gamma_2} - \frac{(\chi^{-1}+\nu^{-1})\delta_2\tau^{-1}}{(\chi^{-1}+\nu^{-1})(\delta^2\chi^{-1}+\delta_1^2\nu^{-1}+\delta_2^2\tau^{-1})-(\delta\chi^{-1}+\delta_1\nu^{-1})^2} \\
&= (1-\lambda)\frac{1-\gamma_2}{\gamma_2} \\
&- \frac{\frac{\chi+\nu}{\chi\nu}\gamma_2\tau^{-1}}{\frac{\chi+\nu}{\chi\nu}\frac{1}{(1-\lambda)+\lambda\gamma_2}(\Xi^2\chi^{-1}+(1-\lambda)^2\gamma_1^2\nu^{-1}+\gamma_2^2\tau^{-1})-\frac{1}{(1-\lambda)+\lambda\gamma_2}(\Xi\chi^{-1}+(1-\lambda)\gamma_1\nu^{-1})^2}
\end{aligned} \tag{B.36}$$

where $\Xi = (1-\lambda)\gamma_1 + \gamma_2$.

Use the definitions of γ_1 and γ_2

$$\begin{aligned}
\hat{\beta}_1 &= (1-\lambda)\frac{\nu+\chi}{\tau} \\
&- \frac{(\chi+\nu)[(1-\lambda)(\nu+\chi)+\tau]\chi}{(\chi+\nu)([(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)-([(1-\lambda)\nu+\tau]+(1-\lambda)^2\nu\chi)^2\nu} \\
&= (1-\lambda)\frac{\nu+\chi}{\tau} - \frac{(\chi+\nu)[(1-\lambda)(\nu+\chi)+\tau]\chi}{\chi\tau(\tau+\nu+\chi)} \\
&= \frac{-\lambda(\nu+\chi)\chi\tau}{\chi\tau(\tau+\nu+\chi)} \\
&= \frac{-\lambda(\nu+\chi)}{(\tau+\nu+\chi)}
\end{aligned} \tag{B.37}$$

negative as long as $0 < \lambda < 1$.

$$\begin{aligned}
\hat{\beta}_2 &= -\frac{\delta_1}{\delta_2} + \frac{-cov(fr,pi)cov(fr,err)}{var(fr)var(pi)-cov(fr,pi)^2} \\
&= -\frac{\delta_1}{\delta_2} + \frac{(\delta\chi^{-1}+\delta_1\nu^{-1})\delta_2\frac{1}{\tau}}{(\chi^{-1}+\nu^{-1})(\delta^2\chi^{-1}+\delta_1^2\nu^{-1}+\delta_2^2\tau^{-1})-(\delta\chi^{-1}+\delta_1\nu^{-1})^2} \\
&= -(1-\lambda)\frac{\gamma_1}{\gamma_2} \\
&- \frac{[(1-\lambda)\nu+\tau]\chi^{-1}+(1-\lambda)}{\frac{\chi+\nu}{\chi\nu}(\Xi^2\chi^{-1}+(1-\lambda)^2\gamma_1^2\nu^{-1}+\gamma_2^2\tau^{-1})-(\Xi\chi^{-1}+(1-\lambda)\gamma_1\nu^{-1})^2}
\end{aligned} \tag{B.38}$$

use definition of γ_1 and γ_2

$$\begin{aligned}\hat{\beta}_2 &= -(1-\lambda)\frac{\nu}{\tau} \\ &+ \frac{([(1-\lambda)\nu+\tau]+(1-\lambda)\chi)\chi\nu}{(\chi+\nu)([(1-\lambda)\nu+\tau]^2+[(1-\lambda)^2\nu+\tau]\chi)-([(1-\lambda)\nu+\tau]+(1-\lambda)^2\nu\chi)^2\nu} \\ &= -(1-\lambda)\frac{\nu}{\tau}+\frac{([(1-\lambda)\nu+\tau]+(1-\lambda)\chi)\chi\nu}{\chi\tau(\tau+\nu+\chi)} \\ &= \frac{\lambda\nu\chi\tau}{\chi\tau(\tau+\nu+\chi)} \\ &= \frac{\lambda\nu}{(\tau+\nu+\chi)}\end{aligned} \tag{B.39}$$

positive as long as $0 < \lambda < 1$. ■

B.3 Different public signal measure

In addition to our baseline measure in section 2.2, we use the current value of the forecasted series as an additional possible proxy for public signal . In particular, assume that the observable series y agents are asked to forecast depends on a latent unobservable factor x and some noise e. Moreover, agents receive some private noisy signal on it s_t^i.

$$\begin{aligned} y_t &= x_t + e_t \\ x_t &= \rho x_{t-1} + u_t \\ s_t^i &= x_t + \eta_t^i \end{aligned} \tag{B.40}$$

with u_t, e_t and η^i normally distributed with zero mean and $\rho < 1$. The observable contemporaneous y_t is a public noisy signal about the underlying fundamental x_t. This structure is consistent with CG and BGMS econometric specification as long as $\tilde{E}_t[y_{t+h}] = \tilde{E}_t[x_{t+h}]$.

To measure the contemporaneous public signal for financial series, we use the average value of the series in the same quarter up to the survey date, which is the second month of the quarter. On the other hand, macroeconomic series are released with some lag, therefore we use the first release of the previous period value, which is available at the time of the forecast. To capture the surprise component in the public information, we compute the difference between the public signal and individual prior about the signal. In this case $pi_{t,t+h} = y_t - E_{t-1}^i[y_t]$. For macroeconomic variables, we compare contemporaneous release of lagged value with lagged nowcasting. In thi case $pi_{t,t+h} = y_{t-1} - E_{t-1}^i[y_{t-1}]$.

We run regression 2.13 using this different measure of public information. Panel A of Table B.1 reports the panel data regressions at 3 quarters horizon with individual fixed effects and the median from individual regressions. The tables displays

Table B.1: Private and public information: alternative measure of public information

Panel A: 3 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.25	0.08	0.00	-0.18	-0.05	0.12	0.69	-0.13
GDP price index inflation	-0.40	0.04	0.00	-0.40	0.39	0.15	0.01	0.30
Real GDP	-0.10	0.08	0.23	0.06	-0.07	0.08	0.39	-0.10
Consumer Price Index	-0.19	0.08	0.03	-0.14	-0.56	0.28	0.06	-0.52
Industrial production	-0.30	0.14	0.03	-0.35	0.08	0.14	0.57	0.11
Housing Start	-0.09	0.09	0.36	-0.13	0.57	0.13	0.00	0.37
Real Consumption	-0.30	0.12	0.01	-0.25	0.27	0.13	0.06	0.15
Real residential investment	-0.09	0.10	0.39	-0.07	0.57	0.18	0.00	0.48
Real nonresidential investment	0.06	0.14	0.65	0.18	0.20	0.22	0.38	0.14
Real state and local government consumption	-0.53	0.05	0.00	-0.53	0.12	0.10	0.24	0.17
Real federal government consumption	-0.47	0.04	0.00	-0.39	0.28	0.09	0.00	0.19
Unemployment rate	0.26	0.16	0.10	0.18	-0.39	0.25	0.12	-0.44
Three-month Treasury rate	-0.26	0.10	0.02	-0.31	0.93	0.26	0.00	1.30
Ten-year Treasury rate	-0.63	0.05	0.00	-0.64	0.61	0.11	0.00	0.62
AAA Corporate Rate Bond	-0.69	0.04	0.00	-0.78	0.80	0.10	0.00	0.75

Panel B: 2 quarters horizon

	Revision				Public signal			
	β_1	SE	p-value	Median	β_2	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nominal GDP	-0.14	0.09	0.11	-0.10	0.12	0.12	0.34	0.04
GDP price index inflation	-0.41	0.04	0.00	-0.38	0.46	0.12	0.00	0.34
Real GDP	-0.09	0.10	0.41	0.07	0.06	0.09	0.51	-0.03
Consumer Price Index	-0.07	0.14	0.59	-0.12	-0.50	0.34	0.16	-0.54
Industrial production	-0.19	0.17	0.26	-0.15	0.35	0.22	0.12	0.32
Housing Start	0.03	0.06	0.67	-0.04	0.29	0.11	0.01	0.27
Real Consumption	-0.25	0.11	0.02	-0.21	0.21	0.13	0.11	0.14
Real residential investment	-0.09	0.09	0.32	-0.12	0.41	0.14	0.00	0.41
Real nonresidential investment	0.13	0.12	0.28	0.17	-0.02	0.20	0.94	-0.11
Real state and local government consumption	-0.40	0.04	0.00	-0.36	0.20	0.10	0.05	0.25
Real federal government consumption	-0.42	0.05	0.00	-0.33	0.29	0.10	0.00	0.08
Unemployment rate	0.22	0.12	0.06	0.20	-0.30	0.18	0.10	-0.28
Three-month Treasury rate	-0.33	0.14	0.02	-0.43	0.78	0.30	0.01	1.04
Ten-year Treasury rate	-0.80	0.06	0.00	-0.92	0.75	0.12	0.00	0.76
AAA Corporate Rate Bond	-0.77	0.05	0.00	-0.83	0.92	0.07	0.00	0.88

Notes: this table reports the coefficients of regression 2.13 (individual forecast errors on individual revisions and public information). Columns 1 to 3 show coefficient β_1 (forecast revision) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 4 shows the median coefficient of the same regression at the individual level. Columns 5 to 7 show coefficient β_2 (public information) from the panel regression with individual fixed effect, with standard errors and corresponding p-values. Standard errors are robust and clustered by forecaster. Column 8 shows the median coefficient of the same regression at the individual level. Panel A uses forecast at 3 quarters horizon and panel B uses forecast at 2 quarters horizon.

consistent $\beta_{GV,1} < 0$ and $\beta_{GV,2} > 0$ across variables, though less consistently than in table 2.6 in section 2.2. The reason being that the measure of public signal considered here doesn't refer direction to horizon $h = 3$, but to horizon $h = 0$ or even ($h = -1$ for macro variable) and it is therefore less informative about longer horizons. Panel B of Table B.1 reports the same regression using a shorter horizon $h = 2$ and shows that the result are much more consistent and significant.[1] Figure **??** shows the coefficients graphically.

B.4 Empirical evidence with AR2

Table B.2: Motivating evidence: BGMS regressions with 2 lags

	$fr^i_{t+2,t}$				$fr^i_{t+1,t}$			
	$\beta_{BGMS,1}$	SE	p-value	Median	$\beta_{BGMS,2}$	SE	p-value	Median
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	
Nominal GDP	-0.24	0.14	0.10	-0.19	0.15	0.15	0.30	0.11
GDP price index inflation	-0.36	0.09	0.00	-0.38	0.27	0.13	0.04	0.36
Real GDP	-0.08	0.16	0.62	0.24	0.08	0.19	0.69	-0.15
Consumer Price Index	-0.89	0.18	0.00	-1.20	0.70	0.28	0.02	1.05
Industrial production	-0.30	0.19	0.12	-0.21	0.25	0.22	0.27	0.09
Housing Start	-0.26	0.15	0.09	-0.32	0.67	0.17	0.00	0.68
Real Consumption	-0.34	0.19	0.08	-0.35	0.34	0.20	0.10	0.29
Real residential investment	-0.54	0.19	0.01	-0.29	0.83	0.20	0.00	0.62
Real nonresidential investment	0.26	0.31	0.42	0.57	-0.05	0.38	0.91	-0.27
Real state and local government consumption	-0.10	0.13	0.46	-0.20	-0.06	0.16	0.70	0.01
Real federal government consumption	-0.46	0.13	0.00	-0.44	0.34	0.14	0.02	0.25
Unemployment rate	0.14	0.22	0.51	0.00	0.21	0.20	0.29	0.26
Three-month Treasury rate	-0.35	0.12	0.01	-0.58	0.75	0.21	0.00	1.25
Ten-year Treasury rate	-0.97	0.13	0.00	-0.96	0.85	0.16	0.00	0.76
AAA Corporate Rate Bond	-0.68	0.14	0.00	-1.08	0.54	0.20	0.01	0.84

[1] At both horizons forecasts about the Consumer Price Index seem to overreact to this measure of public information instead of underreacting. However, in unreported result we show that if we consider the actual value of GDP deflator as a public signal for consumer inflation (highly correlated with CPI), the forecasts underreact to it.

Table B.3: Motivating evidence: public information regressions with 2 lags

	$fr^i_{t+2,t}$				$fr^i_{t+1,t}$				$pi^i_{t+2,t}$				$pi^i_{t+1,t}$			
Variable	$\beta_{fr,1}$ (1)	SE (2)	p-value (3)	Median (4)	$\beta_{fr,2}$ (5)	SE (6)	p-value (7)	Median (8)	$\beta_{pi,1}$ (9)	SE (10)	p-value (11)	Median (12)	$\beta_{pi,2}$ (13)	SE (14)	p-value (15)	Median (16)
Nominal GDP	-0.71	0.21	0.00	-0.55	1.12	0.14	0.00	1.25	0.38	0.21	0.09	0.16	-0.48	0.17	0.01	-0.67
GDP price index inflation	-0.85	0.13	0.00	-0.82	1.17	0.11	0.00	1.21	0.48	0.18	0.01	0.46	-0.57	0.14	0.00	-0.74
Real GDP	-0.58	0.25	0.02	-0.32	1.25	0.22	0.00	1.12	0.42	0.32	0.19	0.14	-0.84	0.33	0.01	-0.67
Consumer Price Index	-1.55	0.27	0.00	-1.53	1.55	0.26	0.00	1.85	1.27	0.39	0.00	1.43	-1.14	0.34	0.00	-1.61
Industrial production	-0.72	0.25	0.01	-0.66	1.08	0.16	0.00	0.78	0.49	0.26	0.07	0.22	-0.52	0.19	0.01	-0.22
Housing Start	-0.75	0.20	0.00	-0.78	1.03	0.14	0.00	1.03	0.97	0.23	0.00	0.84	-0.69	0.18	0.00	-0.63
Real Consumption	-0.72	0.25	0.01	-0.89	1.02	0.21	0.00	0.89	0.51	0.24	0.04	0.57	-0.40	0.23	0.10	-0.29
Real residential investment	-1.03	0.24	0.00	-0.79	1.45	0.21	0.00	1.47	1.17	0.27	0.00	0.90	-1.18	0.24	0.00	-0.73
Real nonresidential investment	0.08	0.38	0.84	0.26	0.63	0.32	0.06	0.95	-0.01	0.45	0.99	-0.34	-0.18	0.43	0.69	-0.66
Real state and local government consumption	-0.46	0.15	0.00	-0.29	0.95	0.19	0.00	0.98	-0.01	0.18	0.98	-0.15	-0.26	0.21	0.23	-0.32
Real federal government consumption	-1.10	0.16	0.00	-0.90	1.35	0.12	0.00	1.07	0.55	0.17	0.00	0.36	-0.62	0.15	0.00	-0.45
Unemployment rate	-0.15	0.30	0.62	-0.33	0.83	0.21	0.00	0.93	0.43	0.27	0.13	0.65	-0.55	0.20	0.01	-0.74
Three-month Treasury rate	-1.13	0.20	0.00	-1.75	1.64	0.19	0.00	2.06	1.43	0.33	0.00	1.98	-1.47	0.35	0.00	-2.31
Ten-year Treasury rate	-1.72	0.19	0.00	-1.68	1.91	0.17	0.00	1.98	1.39	0.25	0.00	1.45	-1.42	0.23	0.00	-1.53
AAA Corporate Rate Bond	-1.50	0.13	0.00	-1.59	1.64	0.15	0.00	1.47	1.10	0.24	0.00	1.09	-1.03	0.25	0.00	-0.90

B.5 Survey anonimity

The forecast data used in this paper are from the Survey of Professional Forecasters, compiled by the Federal Reserve of Philadelphia. Even if this particular survey is anonymous, we argue that it can nonetheless be affected by strategic incentives as well. In particular, we argue that the survey provided by forecasters to anonymous surveys appear to be the same as the one provided to other non-anonymous survey. This has been noted before in the forecasting literature: "According to industry experts, forecasters often seem to submit to the anonymous surveys the same forecasts they have already prepared for public (i.e. non-anonymous) release. There are two reasons for this. First, it might not be convenient for the forecasters to change their report, unless they have a strict incentive to do so. Second, the forecasters might be concerned that their strategic behavior could be uncovered by the editor of the anonymous survey." (Marinovic et al., 2013)

Two observations support this claim. First, Bordalo et al. (2020) establish fact 1 and 2 in section 2.2 by using both the SPF data and the Blue Chip data, which are not anonymous. They show that the two series provide very similar results, which is in line with the hypothesis of forecasters provided similar forecast to both surveys. Second, in a survey by the European Central Bank supplementary to their Survey of Professional Forecasters, respondents are asked explicitly "When responding to the SPF, what forecast do you provide?". In 2013, more than 80% of the panelists responded "the last available, while in 2008 more than 90% gave the same answer (European Central Bank, 2014). It is also important to note that this is a conservative estimate of agents compiling a new forecast exclusively for the ECB survey, as the new forecast provided might be compiled to be used for other non-anonymous surveys as well.

B.6 Dynamic model with AR(2)

We consider here a dynamic setting with a fundamental AR(2) process

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + u_t$$
$$\begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{B.41}$$
$$\bar{X}_t = A\bar{X}_{t-1} + a \begin{bmatrix} u_t \\ e_t \end{bmatrix}$$

With $u \sim N(0, \nu^{-1})$.

Each agent receive a private signal s_t^i and a public signal g_t

$$\begin{aligned} s_t^i &= x_t + \eta_t^i \\ g_t &= x_t + e_t \end{aligned} \tag{B.42}$$

with $\eta_t^i \sim N(0, \tau^{-1})$, $e_t \sim N(0, \nu^{-1})$. In matrix form

$$V_t^i \equiv \begin{bmatrix} g_t \\ s_t^i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \bar{X}_t + \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \tag{B.43}$$

Honest beliefs Agents form beliefs about x at horizon h: $E_t^i[\bar{X}_{t+h}]$. The honest posterior belief about $\bar{X}$ is given by the Kalman filter

$$E_t^i[\bar{X}_t] = AE_{t-1}^i[\bar{X}_{t-1}] + K(V_t^i - E_{t-1}^i[V_t])$$

With the first line yields the posterior $E_t^i[x_t] \equiv x_{t,t}^i$

$$x_{t,t}^i = x_{t,t-1}^i + K_{1,1}(g_t - x_{t,t-1}^i) + K_{1,2}(s_t^i - x_{t,t-1}^i)$$

where the Kalman gains are

$$\begin{aligned} K_{1,1} &= \frac{\nu}{\Sigma^{-1}+\nu+\tau} \\ K_{1,2} &= \frac{\tau}{\Sigma^{-1}+\nu+\tau} \end{aligned} \tag{B.44}$$

and the posterior forecast error variance

$$\Sigma \equiv E[(x_t - x^i_{t,t-1})(x_t - x^i_{t,t-1})'] \tag{B.45}$$

Strategic interactions As in the previous section, the strategic substitutability in agents objective function leads them to report

$$\begin{aligned} \begin{bmatrix} \hat{x}^i_{t,t} \\ \hat{x}^i_{t-1,t} \end{bmatrix} &= \begin{bmatrix} \frac{1}{1-\lambda}E^i_t[x_t] - \frac{\lambda}{1-\lambda}E^i[\bar{\hat{x}}_{t,t}] \\ \frac{1}{1-\lambda}E^i_t[x_t-1] - \frac{\lambda}{1-\lambda}E^i[\bar{\hat{x}}_{t-1,t}] \end{bmatrix} \\ F^i_t &= \begin{bmatrix} \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} E^i_t \begin{bmatrix} \bar{X}_t \\ F_t \end{bmatrix} \end{aligned} \tag{B.46}$$

where $\hat{x}^i_{t+h,t}$ is the forecast provided by individual i in t about realization in $t+h$, and $\bar{\hat{x}}_{t+h,t} = \int^i \hat{x}^i_{t+h,t}di$ is the average of forecasts provided in t about realization in $t+h$. Define $F_t \equiv \begin{bmatrix} \bar{\hat{x}}_{t,t} \\ \bar{\hat{x}}_{t-1,t} \end{bmatrix}$ and $F^i_t \equiv \begin{bmatrix} \hat{x}^i_{t,t} \\ \hat{x}^i_{t-1,t} \end{bmatrix}$. If $\lambda = 0$, agents report their true beliefs. With $1 > \lambda > 0$, agents not only want to be accurate, but also to stand out with respect to the average forecast.

We average $\hat{x}^i_{t,t}$ across agents and use repeated substitution in B.46 to express the reported average forecast as

$$F_t = -\frac{1}{1-\lambda}\sum_{k=0}^{\infty}\left(\frac{\lambda}{1-\lambda}\right)^k \bar{E}^{(k)}[\bar{X}_t] = \frac{1}{1-\lambda}\bar{E}_t\bar{X}_t - \frac{\lambda}{1-\lambda}F_t \tag{B.47}$$

We guess and verify the law of motion for F_t and the other unobserved state variables. In particular, we conjecture that the state vector evolves according to[2]

$$Z \equiv \begin{bmatrix} \bar{X}_t \\ F_t \\ w_t \end{bmatrix} = MZ_{t-1} + m \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{B.48}$$

Where

$$M = \begin{bmatrix} \underset{2x2}{A} & \underset{2x2}{0} & \underset{2x1}{0} \\ \underset{2x2}{G} & \underset{2x2}{L} & \underset{2x1}{0} \\ \underset{1x2}{0} & \underset{1x2}{0} & \underset{1x1}{0} \end{bmatrix} \quad \text{and} \quad m = \begin{bmatrix} \underset{2x2}{a} \\ \underset{2x2}{\mu} \\ 0 \quad 1 \end{bmatrix} \tag{B.49}$$

the observable variables are the two signals about x_t

$$V_t^i \equiv \begin{bmatrix} g_t \\ s_t^i \end{bmatrix} = HZ_t + \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \tag{B.50}$$

where

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{B.51}$$

Agents use their conjecture law of motion B.48 and the observables B.50 to infer the state using the individual Kalman filter. The posterior estimate of the state vector by agent i is

$$\begin{aligned} E_t^i[Z_t] &= ME_{t-1}^i[Z_{t-1}] + K(V_t^i - E_{t-1}^i[V_t]) \\ &= (I - KH)ME_{t-1}^i[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} + K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \end{aligned} \tag{B.52}$$

[2] w_t takes care of the correlation between public signal and higher order beliefs F_t

Where K is the Kalman gain. Average B.52 to find the consensus believe on the state vector.

$$\bar{E}_t[Z_t] = (I - KH)M\bar{E}_{t-1}[Z_{t-1}] + KHMZ_{t-1} + KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \tag{B.53}$$

From the definition on F_t in 2.28 it follows that

$$\begin{aligned} F_t &= \begin{bmatrix} \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} & 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} \bar{E}_t[Z_t] \equiv \xi \bar{E}_t[Z_t] \\ &= \xi(I - KH)M\bar{E}_{t-1}[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \end{aligned} \tag{B.54}$$

Compute (i) $\xi M\bar{E}_{t-1}[Z_{t-1}]$, (ii) $HM\bar{E}_{t-1}[Z_{t-1}]$, (iii) HMZ_{t-1}, (iv) Hm.

1. write ξ as a vector of matrices

$$\xi \equiv \begin{bmatrix} \begin{bmatrix} \frac{1}{1-\lambda} & 0 \\ 0 & \frac{1}{1-\lambda} \end{bmatrix} & \begin{bmatrix} -\frac{\lambda}{1-\lambda} & 0 \\ 0 & -\frac{\lambda}{1-\lambda} \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{bmatrix} \equiv \begin{bmatrix} \frac{1}{1-\lambda}I & -\frac{\lambda}{1-\lambda}I & 0 \end{bmatrix} \tag{B.55}$$

Then

$$\begin{aligned} \xi M\bar{E}_{t-1}[Z_{t-1}] &= \begin{bmatrix} \frac{1}{1-\lambda}I & -\frac{\lambda}{1-\lambda}I & 0 \end{bmatrix} \begin{bmatrix} A & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \bar{E}_{t-1}[Z_{t-1}] \\ &= \begin{bmatrix} \frac{1}{1-\lambda}A - \frac{\lambda}{1-\lambda}G & -\frac{\lambda}{1-\lambda}L & 0 \end{bmatrix} \bar{E}_{t-1} \begin{bmatrix} \bar{X}_{t-1} \\ F_{t-1} \\ w_{t-1} \end{bmatrix} \\ &= \left(\frac{1}{1-\lambda}A - \frac{\lambda}{1-\lambda}G\right) \bar{E}_{t-1}[\bar{X}_{t-1}] - \frac{\lambda}{1-\lambda}L\bar{E}_{t-1}[F_{t-1}] \end{aligned} \tag{B.56}$$

2. write H as a vector of matrices

$$H \equiv \begin{bmatrix} \begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{1,1} & H_{1,2} \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} \equiv [H_1 \quad H_2 \quad H_3] \tag{B.57}$$

Then

$$\begin{aligned} HM\bar{E}_{t-1}[Z_{t-1}] &= [H_1 \quad H_2 \quad H_3] \begin{bmatrix} A & 0 & 0 \\ G & L & 0 \\ 0 & 0 & 0 \end{bmatrix} \bar{E}_{t-1}[Z_{t-1}] \\ &= [H_1 A \quad 0 \quad 0] \bar{E}_{t-1} \begin{bmatrix} \bar{X}_{t-1} \\ F_{t-1} \\ w_{t-1} \end{bmatrix} \\ &= H_1 A \bar{E}_{t-1}[\bar{X}_{t-1}] \end{aligned} \tag{B.58}$$

3. Similarly,

$$HMZ_{t-1} = H_1 A \bar{X}_{t-1} \tag{B.59}$$

4. Similarly

$$Hm = H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{B.60}$$

Substitute back in the posted KF, using that $-\frac{\lambda}{1-\lambda}\bar{E}_{t-1}[F_{t-1}] = F_{t-1} - \frac{1}{1-\lambda}\bar{E}_{t-1}[X_{t-1}]$.

After some algebra, one gets

$$\begin{aligned} F_t = & \left(\frac{1}{1-\lambda}A + -\frac{\lambda}{1-\lambda}G - \frac{1}{1-\lambda}L - \xi K H_1 A\right)\bar{E}_{t-1}[\bar{X}_{t-1}] \\ & + \xi K H_1 A \bar{X}_{t-1} + LF_{t-1} + \xi K \left(H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} u_t \\ e_t \end{bmatrix} \end{aligned} \tag{B.61}$$

Equation B.61 must equal the second line (a 2x1 vector) of the perceived law of motion B.48. The solution to the fixed point is given by $G = \xi K H_1 A$, $\mu = \xi K \left(H_1 a + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)$ and $L = A - G$.

In particular, define

$$\begin{aligned} C_1 \equiv \frac{K_{1,1} - \lambda(K_{3,1})}{1-\lambda}, \qquad C_2 \equiv \frac{K_{1,2} - \lambda K_{3,2}}{1-\lambda} \qquad \text{and} \qquad C = C_1 + C_2 \\ D_1 \equiv \frac{K_{2,1} - \lambda(K_{4,1})}{1-\lambda}, \qquad D_2 \equiv \frac{K_{2,2} - \lambda K_{4,2}}{1-\lambda} \qquad \text{and} \qquad D = D_1 + D_2 \end{aligned}$$

Then $G = \begin{bmatrix} \rho_1 C & \rho_2 C \\ \rho_1 D & \rho_2 D \end{bmatrix}$, $\mu = \begin{bmatrix} C & C_1 \\ D & 0 \end{bmatrix}$ and $L = \begin{bmatrix} \rho_1(1-C) & \rho_2(1-C) \\ 1-\rho_1 D & -\rho_2 D \end{bmatrix}$.

Given the law of motion of unobserved state 2.29 and the observable 2.31, the posterior variance of the forecast solves the following Ricatti equation

$$\begin{aligned} \Sigma & \equiv E[(Z_t - Z^i_{t,t-1})(Z_t - Z^i_{t,t-1})'] \\ \Sigma & = M(\Sigma - \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix}\right)^{-1} H\Sigma)M' + m \begin{bmatrix} \xi^{-1} & 0 \\ 0 & \nu^{-1} \end{bmatrix} m' \end{aligned} \tag{B.62}$$

and the Kalman filter is

$$K = \Sigma H' \left(H\Sigma H' + \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix}\right)^{-1} \tag{B.63}$$

Step 6: derive the action of individual With the model's solution, one can obtain the individual forecast as

$$\begin{aligned}
F_t^i &= \xi \quad E_t^i[Z_t] \\
&= \xi(I-KH)ME_{t-1}^i[Z_{t-1}] + \xi KHMZ_{t-1} + \xi KHm \begin{bmatrix} u_t \\ e_t \end{bmatrix} \\
=& A \underbrace{\left(\frac{1}{1-\lambda}E_{t-1}^i[\bar{X}_{t-1}] - \frac{\lambda}{1-\lambda}E_{t-1}^i[F_{t-1}]\right)}_{F_{t-1}^i} \\
&+ [-\frac{\lambda}{1-\lambda}G - G]E_{t-1}^i[x_{t-1}] + \frac{\lambda}{1-\lambda}GE_{t-1}^i[F_{t-1}] \\
&- \xi KH_1AE_{t-1}^i[\bar{X}_{t-1}] + \xi KH_1A\bar{X}_{t-1} + \xi KH_1a \begin{bmatrix} u_t \\ e_t \end{bmatrix} + \xi K \begin{bmatrix} e_t \\ 0 \end{bmatrix} + \xi K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \\
F_t^i - AF_{t-1}^i &= -\xi KH_1AF_{t-1}^i + \xi KH_1\bar{X}_t + \xi K \begin{bmatrix} e_t \\ 0 \end{bmatrix} + \xi K \begin{bmatrix} 0 \\ \eta_t^i \end{bmatrix} \\
F_t^i - AF_{t-1}^i &= \xi K \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} (\bar{X}_t - AF_{t-1}^i) + \xi K \begin{bmatrix} 1 \\ 0 \end{bmatrix} e_t + \xi K \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_t^i \\
F_t^i - AF_{t-1}^i &= \begin{bmatrix} C \\ D \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} (\bar{X}_t - AF_{t-1}^i) + \begin{bmatrix} C_1 \\ D_1 \end{bmatrix} e_t + \begin{bmatrix} C_2 \\ D_2 \end{bmatrix} \eta_t^i
\end{aligned} \tag{B.64}$$

consider the first line

$$\hat{x}_{t,t}^i - \hat{x}_{t,t-1}^i = C[x_t - \hat{x}_{t,t-1}^i] + C_1e_t + C_2\eta_t^i \tag{B.65}$$

Which is similar to the basic framework in section 2.2. Consider the second line

$$\hat{x}_{t-1,t}^i - \hat{x}_{t-1,t-1}^i = D[x_{t-1} - \hat{x}_{t-1,t-1}^i] + D_1e_t + D_2\eta_t^i \tag{B.66}$$

B.7 Structural estimation at: 2 quarters horizon

Table B.4: Estimasted parameters

Variable	ρ (1)	$\frac{\sigma_e}{\sigma_u}$ (2)	$\frac{\sigma_\eta}{\sigma_u}$ (3)	λ (4)
Nominal GDP	0.93	1.51	1.31	0.61
GDP price index inflation	0.90	1.10	1.08	0.32
Real GDP	0.80	1.20	1.19	0.38
Consumer Price Index	0.97	1.14	1.22	0.56
Industrial production	0.85	1.41	1.16	0.29
Housing Start	0.85	2.12	1.20	0.31
Real Consumption	0.73	1.05	1.32	0.39
Real residential investment	0.89	1.44	1.20	0.23
Real nonresidential investment	0.88	3.16	1.04	0.14
Real state and local government consumption	0.74	1.07	1.67	0.73
Real federal government consumption	0.77	1.11	1.61	0.69
Unemployment rate	0.97	3.15	1.03	-0.28
Three-month Treasury rate	0.94	3.16	1.03	0.06
Ten-year Treasury rate	0.83	1.48	1.39	0.69
AAA Corporate Rate Bond	0.85	2.13	1.53	0.83

Table B.5: Moments in data and model

	Targeted moments						Untargeted moments					
	Mean Dispersion		C		β_1		β_{CG}		β_{BGMS}		β_2	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Nominal GDP	0.94	0.94	0.61	0.61	-0.35	-0.35	0.20	0.43	-0.11	-0.22	0.62	0.11
GDP price index inflation	0.34	0.34	0.70	0.70	-0.20	-0.20	0.48	0.20	-0.02	-0.06	0.44	0.10
Real GDP	0.69	0.69	0.63	0.63	-0.25	-0.25	0.33	0.28	-0.07	-0.10	0.54	0.11
Consumer Price Index	0.27	0.27	0.70	0.70	-0.38	-0.38	-0.05	0.21	-0.24	-0.14	0.51	0.20
Industrial production	2.49	2.49	0.59	0.59	-0.16	-0.16	0.33	0.41	-0.01	-0.09	0.49	0.05
Housing Start	75.76	75.76	0.53	0.53	-0.15	-0.15	0.91	0.75	0.12	-0.13	0.54	0.01
Real Consumption	0.34	0.34	0.63	0.63	-0.31	-0.31	0.12	0.16	-0.11	-0.07	0.61	0.16
Real residential investment	16.69	16.69	0.56	0.56	-0.13	-0.13	0.56	0.42	0.07	-0.07	0.49	0.04
Real nonresidential investment	5.02	5.02	0.61	0.59	-0.02	-0.05	0.53	0.67	0.10	-0.05	0.41	0.00
Real state and local government consumption	0.92	0.92	0.61	0.61	-0.65	-0.65	0.05	0.15	-0.24	-0.23	0.77	0.35
Real federal government consumption	4.40	4.40	0.60	0.60	-0.60	-0.60	-0.21	0.18	-0.27	-0.21	0.77	0.31
Unemployment rate	0.09	0.06	0.56	0.55	0.09	0.08	0.59	0.78	0.20	0.08	0.39	0.00
Three-month Treasury rate	0.21	0.12	0.63	0.60	0.02	-0.02	0.40	0.66	0.14	-0.02	0.48	0.00
Ten-year Treasury rate	0.12	0.12	0.60	0.60	-0.46	-0.46	-0.09	0.45	-0.24	-0.31	0.71	0.14
AAA Corporate Rate Bond	0.25	0.25	0.61	0.61	-0.49	-0.49	0.05	0.58	-0.22	-0.44	0.70	0.07

Table B.6: Posted and honest moments

	Gain			Consensus MSE			Dispersion		
	Posted	Honest	Ratio	Posted	Honest	Ratio	Posted	Honest	Ratio
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Nominal GDP	0.61	0.49	0.80	0.27	0.60	2.22	0.94	0.41	0.44
GDP price index inflation	0.70	0.66	0.94	0.27	0.37	1.36	0.34	0.22	0.65
Real GDP	0.63	0.57	0.91	0.56	0.80	1.41	0.69	0.39	0.57
Consumer Price Index	0.70	0.62	0.89	0.13	0.23	1.86	0.27	0.10	0.39
Industrial production	0.59	0.54	0.92	1.71	2.29	1.34	2.49	1.79	0.72
Housing Start	0.53	0.47	0.87	35.99	50.84	1.41	75.76	57.87	0.76
Real Consumption	0.63	0.59	0.95	0.57	0.71	1.24	0.34	0.16	0.48
Real residential investment	0.56	0.53	0.94	13.81	17.09	1.24	16.69	12.95	0.78
Real nonresidential investment	0.59	0.57	0.95	2.08	2.43	1.17	5.02	4.65	0.93
Real state and local government consumption	0.61	0.55	0.89	0.73	1.10	1.51	0.92	0.11	0.12
Real federal government consumption	0.60	0.53	0.88	3.60	5.48	1.52	4.40	0.69	0.16
Unemployment rate	0.55	0.60	1.08	0.04	0.03	0.78	0.06	0.07	1.12
Three-month Treasury rate	0.60	0.59	0.98	0.05	0.05	1.07	0.12	0.11	0.97
Ten-year Treasury rate	0.60	0.44	0.73	0.03	0.08	2.48	0.12	0.04	0.29
AAA Corporate Rate Bond	0.61	0.31	0.51	0.02	0.10	4.58	0.25	0.06	0.24

Appendix C

International Trade and Portfolio Diversification: the Role of Information

C.1 Appendix

C.1.1 Fixed point problem

Substituting optimal portfolios 3.14 for domestic and foreign agent in the equilibrium condition 3.15 one can find the price for asset $k \in \{h, f\}$

$$p_k = \frac{1}{R}\bar{\sigma}_k^2 \left[\bar{E}(f_k) - \gamma(\bar{z}_k + z_k) - \gamma\frac{1}{2}\left(\alpha c_k V_k + (1-\alpha)c_k V_k^*\right)\right] \tag{C.1}$$

where

$$\bar{\sigma}_k^2 = \left[\frac{1}{2}\left(\frac{1}{b_k^2\hat{\sigma}_k^2 + \sigma_{fk}^2} + \frac{1}{b_k^2\hat{\sigma}_k^{2*} + \sigma_{fk}^2}\right)\right]^{-1} \tag{C.2}$$

is the average posterior variance, and

$$\bar{E}(f_k) = \left[\frac{1}{2} \left(\frac{1}{b_k^2 \hat{\sigma}_k^2 + \sigma_{fk}^2} \int^H (\bar{f}_k + b_k \hat{M}_k^{(i)}) di + \frac{1}{b_k^2 \hat{\sigma}_k^{2*} + \sigma_{fk}^2} \int^F (\bar{f}_k + b_k \hat{M}_k^{(i)*}) di \right) \right] \tag{C.3}$$

is the average posterior mean. For notational convenience, define

$$\begin{aligned} V_k &= \frac{b_k \hat{\sigma}_k^2}{b_k^2 \hat{\sigma}_k^2 + \sigma_{fk}^2} \\ V_k^* &= \frac{b_k \hat{\sigma}_k^2}{b_k^2 \hat{\sigma}_k^2 + \sigma_{fk}^2} \end{aligned} \tag{C.4}$$

and

$$\bar{q}_k = \left(V \frac{1}{2\sigma_{\eta k}^2} + V^* \frac{1}{2\sigma_{\eta k}^{2*}} \right) \tag{C.5}$$

The solution to the fixed point problem is of the form (3.15) with

$$\begin{aligned} \bar{\lambda}_k &= \frac{1}{R} \bar{\sigma}_k^2 \left[\frac{\bar{f}_k}{\bar{\sigma}_k^2} - \gamma \bar{z}_k - \gamma \frac{1}{2} (\alpha c_k V_k + (1-\alpha) c_k V_k^*) \right] \\ \lambda_{zk} &= -\frac{1}{\gamma R} \bar{\sigma}_k^2 \left(1 + \frac{\bar{q}_k}{\gamma^2 \sigma_z^2} \frac{1}{2} (V + V^*) \right) \\ \lambda_{Mk} &= \frac{1}{R} \bar{\sigma}_k^2 \bar{q}_k \left(1 + \frac{\bar{q}_k}{\gamma^2 \sigma_z^2} \frac{1}{2} (V + V^*) \right) \end{aligned} \tag{C.6}$$

C.1.2 First stage problem

The coefficients A_k and B_k for $k \in \{h, f\}$ in 3.19 are the following:

$$\begin{aligned} A_k &= b_k^2 \sigma_k^2 + R^2 \lambda_{M_k}^2 \sigma_k^2 + R^2 \lambda_{z_k}^2 \sigma_z^2 - 2 b_k R \lambda_{M_k} \sigma_k^2 \\ B_k &= \bar{\sigma}_k^2 \bar{z}_k \gamma + \bar{\sigma}_k^2 \gamma \frac{1}{2} (\alpha c_k V + (1-\alpha) c_k V^*) \end{aligned} \tag{C.7}$$

These two terms are taken as given in the attention allocation problem by each agents, but in equilibrium they depend and affect their choices.

C.1.3 Alternative parametrizations

Asset return and production negatively correlated

Suppose country risk factors affect positively non-financial income but negatively asset's return ($c_k > 0$ and $b_k < 0$). Optimal portfolios present a hedging term with opposite sign as in the previous case, but attention can still decrease non-financial income perceived volatility.

Even without information choice, the risk hedging term by itself is able to yield the prediction that higher trade openness leads to higher portfolio diversification. However, one would need this risk hedging term to be too large to justify the amount of home bias in data. My model provide an endogenous mechanism to amplify portfolio specialization in this case. Nevertheless, the main intuition is still valid: higher the amount of trade, lower the portfolio home bias.

Calibration discussion

The results showed in the previous section relies on a particular calibration. In particular: (P1) in the no trade case ($\alpha = 1$) the benefit from information specialization has to offset the risk hedging term; (P2) in the full trade case ($\alpha = 0.5$) the consumption volatility due to trade has to make the information problem less convex to avoid corner solutions. I now discuss how each parameter is involved in these issues.

First, the "unlearnable variance" σ^2_{fh}: higher is this term in the asset return variance, lower is the increasing return to information. The intuition is that the signal agents pay attention to becomes less informative about the total asset return variance. Therefore higher is this term, lower is the return to information. If it is too high, information specialization in the no trade case is not enough to offset risk hedging (P1). If it is too low, there is information (and portfolio) specialization even with trade (P2).

Second, the correlation between country risk factor and endowment shock c: it drives the benefit of information on consumption volatility, but it also increases the portfolio risk hedging term. If it is too high, the model still predict specialization in attention but domestic portfolio is biased to foreign asset because of risk hedging (P1). If it is too low, the model is not able to yield information (and portfolio) diversification with full trade (P2).

Third, the correlation between asset return and country risk factor b: it increases the benefit of information on asset return (converse of σ^2_{fh}), but it also increases the portfolio risk hedging term (similar to c). If it is too low,the benefit from information is not enough to offset risk hedging in the no trade case (P1). If it is too high, there is information (and portfolio) specialization even with trade (P2).

Moreover, the amount of attention available κ and priors σ_k also affect the problem's convexity in information.

Finally, in the baseline calibration attention allocation does not reach a corner solution for values lower than $\alpha = 1$. If it happens, increasing further α does not increase the investment home bias, since it is not possible to increase attention to domestic factor. Conversely, there is an increase in the hedging term, and therefore a increase in the holding of foreign asset.

Figure C.1 and C.2 show information and portfolio allocations under a calibration leading to the first problem: information allocation is not enough to have portfolio home bias in no trade case ($\alpha = 1$). It embed the case of high σ^2_{fh}, high c_k and low b_k.

Figure C.3 and C.4 show information and portfolio allocations under a calibration leading to the second problem: the information problem is too convex and there is no information diversification in trade case ($\alpha = 0.5$). It embed the case of low σ^2_{fh}, low c_k and high b_k.

Figure C.5 and C.6 show information and portfolio allocations under a calibration leading to the third problem: information allocation is completely specialized for $\alpha < 1$. Therefore, higher α does not increase attention to domestic asset: it increases only risk hedging term. From then on, the share of domestic asset in home agent portfolio decreases instead of increasing.

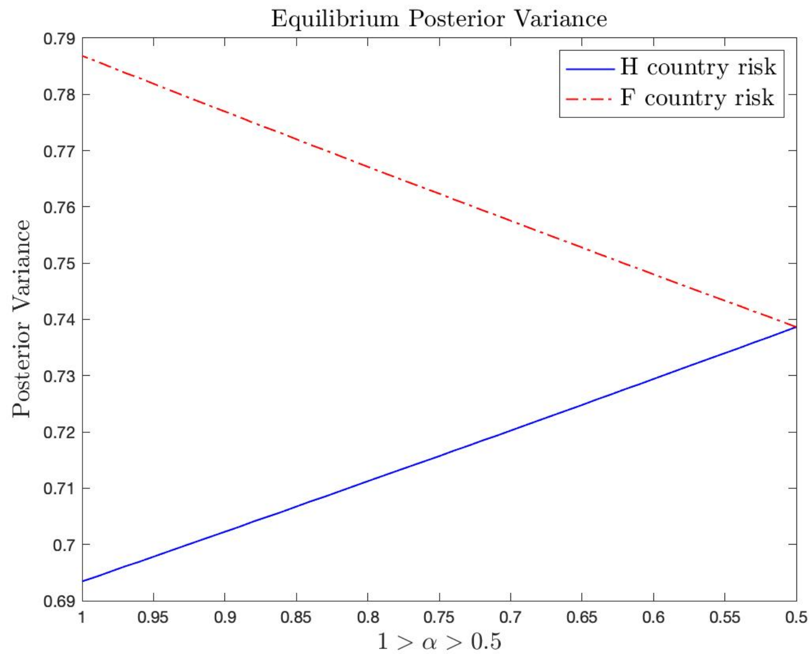


Figure C.1: Equilibrium posterior variance: P1 case (high σ^2_{fh})

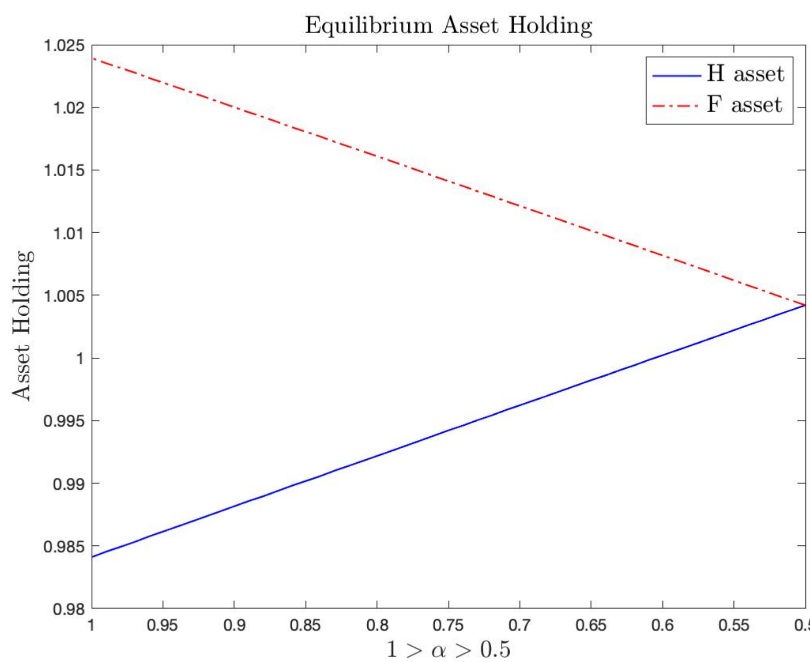


Figure C.2: Equilibrium portfolio allocation: P1 case (high σ^2_{fh})

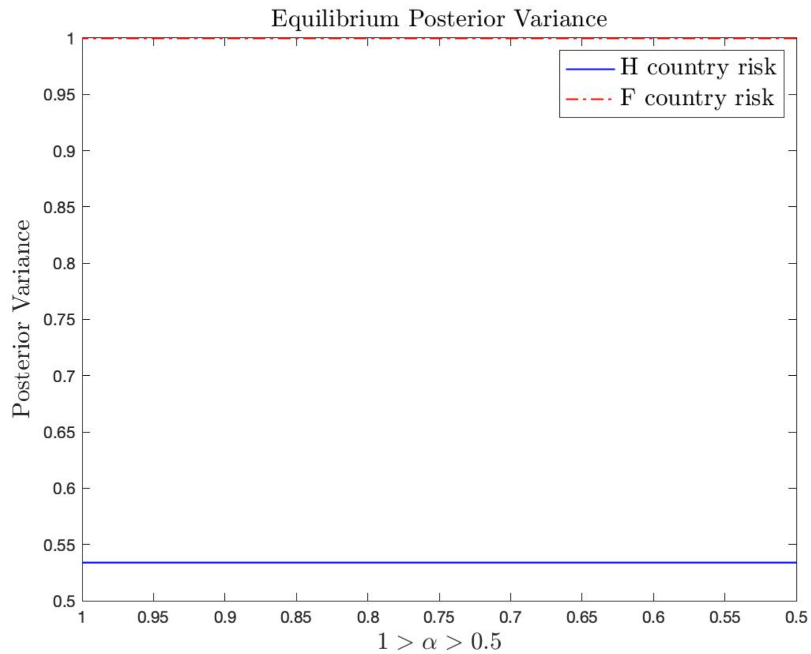


Figure C.3: Equilibrium posterior variance: P1 case (low σ_{fh}^2)

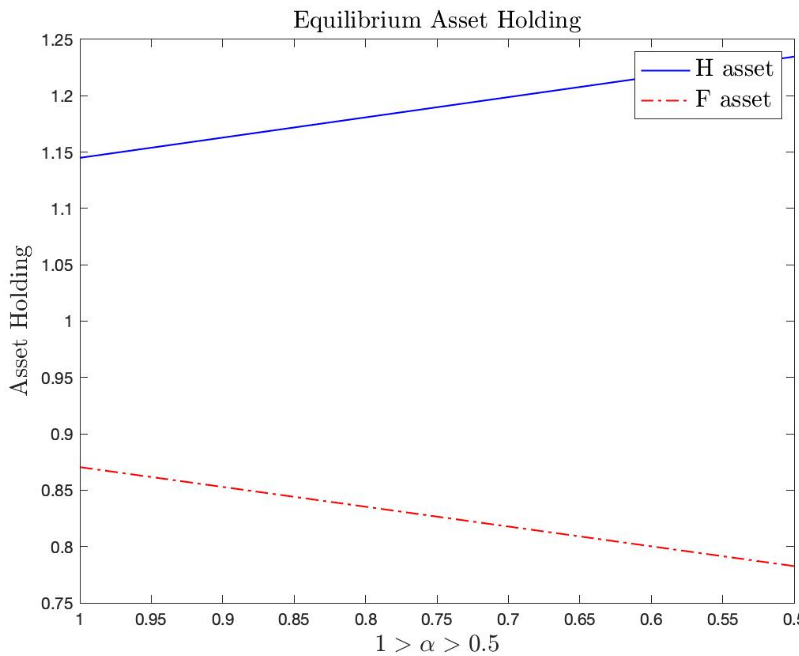


Figure C.4: Equilibrium portfolio allocation: P1 case (low σ_{fh}^2)

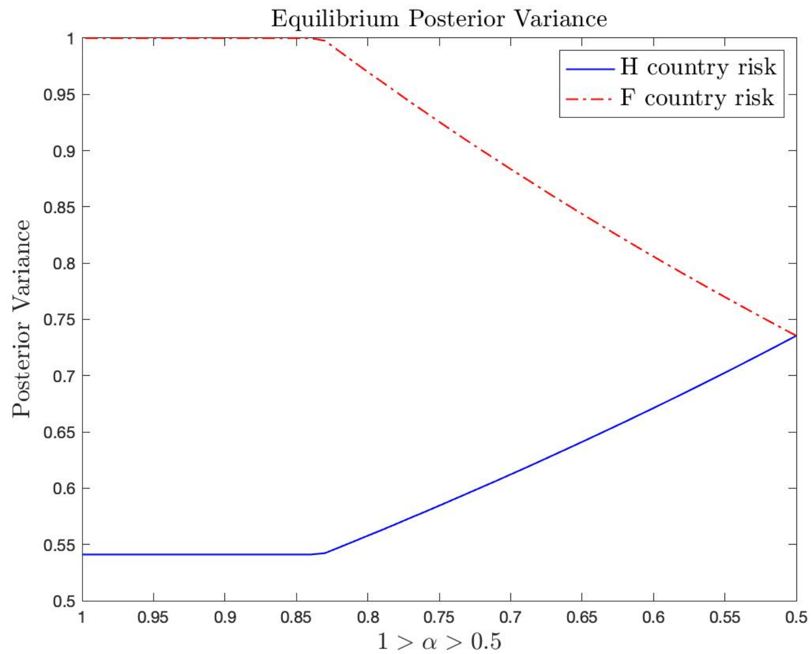


Figure C.5: Equilibrium posterior variance with boundary cases

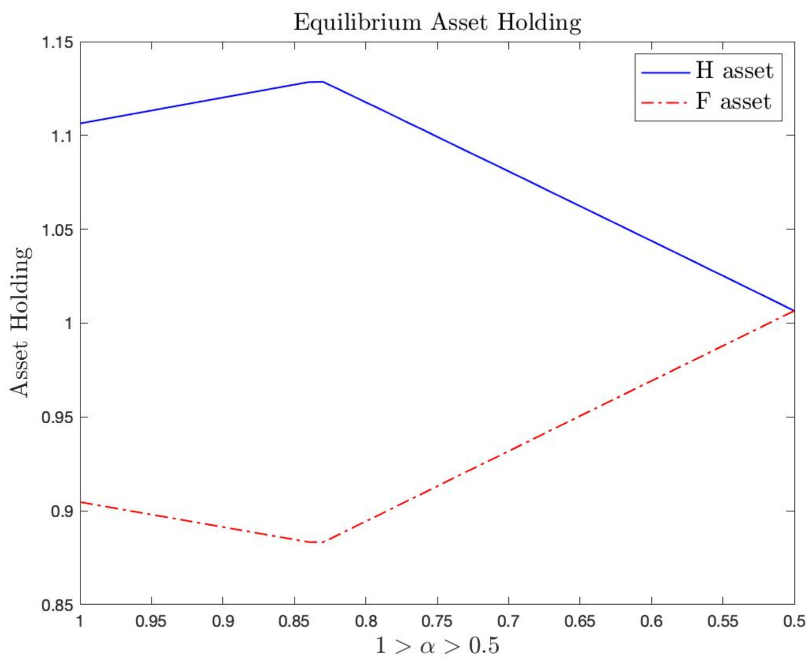


Figure C.6: Equilibrium portfolio allocation with boundary cases