The People versus the Markets: A Parsimonious Model of Inflation Expectations*

Ricardo Reis

LSE

November 2020

Latest version here

Abstract

Expected long-run inflation is sometimes inferred using market prices, other times using surveys. The discrepancy between the two measures has large business-cycle fluctuations, is systematically correlated with monetary policies, and is mostly driven by disagreement between households and traders, and between different traders. A parsimonious model that captures both the dispersed expectations in surveys, and the trading of inflation risk in financial markets, can make sense of the data, and it provides estimates of the underlying expected inflation anchor. Applied to US data, the estimates suggest that inflation became gradually, but steadily, unanchored from 2014 onwards. The model detects this from the fall in cross-person expectations skewness, first across traders, then across people. In general equilibrium, when inflation and the discrepancy are jointly determined, monetary policy faces a trade-off in how strongly to respond to the discrepancy.

*Contact: r.a.reis@lse.ac.uk. I am grateful to comments from Gianluca Benigno, Rodrigo Guimaraes, Peter Kondor, Ian Martin, Carolin Pflueger, Harald Uhlig, and to seminar participants at Chicago Booth, Columbia University, FRB Cleveland, LSE, Oxford University, University of Essex, University of Maryland, the 2020 EABCN - Banque de France conference, the 2019 EABCN - National Bank of Poland conference, the 50th Konstanz seminar, the 2019 MacCalm conference, and the 2019 Rebuilding Macroeconomics conference, and to Jose Alberto Ferreira, Maxi Guennewig, Chao He and, especially, Adrien Couturier for research assistance. This paper was finished while the author was a Lamfalussy Fellow at the Bank for International Settlements. This project has received funding from the European Union's Horizon 2020 research and innovation programme, INFL, under grant number No. GA: 682288.

1 Introduction

Expectations are the bedrock of dynamic economic models. Among them, inflation expectations attract special attention. Inflation is the most commonly asked variable to forecast in surveys, and the central questions in monetary economics—What explains inflation? How can central banks control it? What is the trade-off between inflation and real activity?—depend crucially on what private agents expect inflation will be. Most academic attention has fallen on short-term (usually one-year ahead) inflation expectations by households and firms. Yet, every central banker instead repeats that her focus is on anchoring long-term expectations, and these are usually measured using market prices.

This difference matters. At the end of 2010, with the global financial crisis behind but the Eurozone (EZ) sovereign debt crisis raging, the median short-term EZ inflation expectation was 1.5%, the long-term median was 1.9%, and the market long-term forecast was 1.65%.[1] Four years later, the survey long-term forecast was almost the same, at 1.8%, but the market one had plunged to 0.71%. Policymakers focussed on the latter to justify the introduction of quantitative easing in the EZ, while stating that the former was the crucial one for success.[2] Since the start of 2019, long-term market-based measures have been persistently falling, in both the US and the EZ, while long-term survey-based measures have barely changed. Are inflation expectations anchored or not? Are extra unconventional monetary policies justified? Why do markets and people differ in the first place and what does this tell us about how expectations are formed and how policy affects the economy?

This paper makes three contributions to answering these questions. First, it proposes a new object for study: the business-cycle dynamics of the discrepancy between market and survey measures of long-run inflation expectations; for short, the *discrepancy*. This variable combines three characteristics. First, it concentrates on long-horizon inflation expectations, averaged over 5 or 10 years out. Second, it measures their fluctuations at a business-cycle frequency. Third, it focuses on why two legitimate ways to measure expectations, surveys of people, or prices of assets, give very different estimates. These three characteristics make this variable quite useful for economists trying to understand,

[1]The survey measures are from the Survey of Professional Forecasters, at the 1 and 5 year horizon respectively, and the market measure is from inflation swaps at a 5-year horizon

[2]Reporting on the ECB's president, Mario Draghi's, 2014 speech at Jackson Hole which pointed to the 5-year-5-year forward market expectation fo inflation, the Financial Times wrote "Mr Draghi had highlighted the inflation swap rate ... never before August's Jackson Hole speech had a president of the ECB made such a clear link between its behavior and policy action." But in 2018, in Sintra, reflecting on the unconventional policy measures, Draghi showed the 5-year survey expected inflation to conclude success because:"What is key is that inflation expectations remain well anchored".

in general, how expectations are formed. The focus on the long-horizon is useful from the perspective of behavioral theories of beliefs since differences in (i) when the expectations are measured, (ii) the timing of data releases on present inflation, (iii) anchoring effects on surveys, and (iv) asymmetric payoffs in financial prices, can all have a large impact on 1-year ahead expectations, and complicate the study of expectations alone, but are less relevant for long-run expectations. The focus on business-cycle frequencies raises macroeconomic questions, as opposed to the financial questions in low and high frequency studies of inflation risk premium and liquidity, respectively. Finally, the focus on the difference between market prices and surveys creates a challenge that typical papers on expectations do not meet, since it disciplines the model of subjective beliefs by not just the observed survey responses but also by the equilibrium prices that result.

The second novelty is an empirically-grounded, structurally-justified, *parsimonious model of expectations* for both people and markets that is flexible enough that it can be used as a measurement tool. An active literature over the last two decades has explored different behavioral theories of expectations and different models of market prices, testing some of their main qualitative predictions using surveys and other data. Instead of proposing a new theory of expectations, or a new mechanism for why information is not incorporated, this paper takes a different approach by writing a model that builds on some of the leading theories, and fits the data exactly, with a minimal number of free parameters. I call this a parsimonious model. The model's inputs are asset prices and cross-sectional moments from surveys, and it provides both a decomposition of the discrepancy, as well as a measure of the underlying rational expectation of inflation that is the anchor for inflation dynamics. Crucial for the estimates turns out to be the third moment of the distribution of expectations across traders and people, the skewness.

The third and final part of the paper integrates the model of expectations into a simple, standard, general-equilibrium model of inflation with interest-rate rules. The discrepancy affects inflation by changing both the effective real interest in savings decisions, as well as the nominal interest rate that the central bank chooses to set. The joint endogeneity of inflation and the discrepancy requires a more aggressive policy response to inflation, implies a trade-off when considering how much to respond to markets, and justifies why ex post people often end up forecasting as well as markets.

Outline and findings. Section 2 presents data to measure the discrepancy and establishes three facts about it. First, it has a large business-cycle component, robust across different measures. Second, it is systematically correlated with the state of monetary policy, and is significantly affected by policy shocks and policy regimes. Third, the discrepancy can be decomposed into three terms. The first is compensation for inflation risk required by a fictitious representative agent. The second is a difference between the subjective expectations of the public and that of market traders. The third is the difference between the subjective expectation of the marginal and the average trader. US data suggests that the two last terms drive most of the discrepancy. The section concludes by clarifying that the discrepancy is related, but is not the same, as an inflation risk premium.

Section 3 presents a parsimonious model of how people form subjective expectations that tries to capture some of the main insights from the literatures on imperfect information, over-confidence, learning from experience, and sticky information. The model is sufficiently flexible to exactly fit the three non-zero cross-sectional moments in the distribution of expectations in the Michigan survey of households. Through the lenses of this model, the data reveals an interesting pattern: since around 2010, and more markedly after 2015, the dispersion of household survey expectations has been steadily falling, and at the same time, there has been a steady decrease in the positive skewness of expectations that arises from positive-biases due to experience.

Section 4 inserts these agents in an equilibrium model of dispersed information and financial markets. In the spirit of Grossman and Stiglitz (1980), financial participants can observe market prices, which allows them to form more accurate forecasts than households. However, prices are contaminated by "noise" in the form a supply shock that can be interpreted as resulting from noise traders, animal spirits, or market fads. The novel focus is again on the cross-section of expectations across traders, and especially on the difference between the marginal and the average trader. The model's output exactly matches the decomposition of the discrepancy from section 2.

Section 5 applies the model to US data. The model takes as inputs the cross-sectional moments of surveys of expectations for households, the median expectations from a survey of traders, and the observed asset prices. It fits the discrepancy perfectly, and produces two outputs. The first is the decomposition of the discrepancy between the two disagreement terms. The second is an estimate of the fundamental rational-expectation best forecast of 5-year ahead inflation. The main finding is that long-run inflation expectations have not been anchored in the US. In particular, they have been steadily declining from 2.0% to around 1.8% between 2014 and 2019.

The model provides the following account of why this happened given the observables. In 2014-16, market-expected inflation fell significantly. During that time, people's survey expected inflation stayed steady, as did the median bond-trader survey expected inflation. The model attributes this large discrepancy to the marginal trader in the market expecting much lower inflation, while the median trader is steady, and so to a negative skew in the traders' cross-sectional distribution. This fall in the marginal trader's expectation is mostly due to a noise financial shock, but partly also due to a fall in the fundamental expectations anchor. Then, after 2016, the bond trader's whole distribution moves left, with the median and the marginal trader converging, so the skew goes back to zero. This confirms the negative shock to the fundamental. Right after 2016, it is the disagreement across agents, of traders versus people, that drives the model estimates of a falling fundamental expected inflation Finally, throughout this time, there is a steady decline in the dispersion of household survey expectations, as well as in their positive skew. The model interprets this as people slowly converging towards the lower expectations of traders. It confirms the estimates that fundamental rational expectations of long-run inflation have persistently fallen.

Section 6 inserts the model of expectations and financial markets into a model of monetary policy and inflation. The discrepancy affects monetary policy through two channels. First, because it provides a signal of fundamental expected inflation. Second, because it can transmit noise from financial markets into realized inflation volatility through the transmission of shocks to outcomes and monetary policy's response to noisy indicator of expectations. The section shows how a non-zero discrepancy affects the conditions for a determinate equilibrium for inflation, the sensitivity of inflation expectation to shocks, and the variability of inflation. For all three, the equilibrium feedback between the discrepancy and inflation is important.

Finally, section 7 concludes.

Link to the literature. A long literature has studied inflation expectations in economics. Most of that literature focuses on inflation over the next year though.[3] The focus of this paper is instead on long-run inflation expectations, and the measures used are of inflation on average over the next 5 or 10 years.[4] Another significant literature has focussed on disagreement on inflation expectations. Some focus on the second moment within surveys,

[3]For recent surveys, see Coibion, Gorodnichenko and Kamdar (2018) applied to the Phillips curve, Clements (2019) applied to forecasting, and Kose et al. (2019) applied to emerging markets.

[4]Other papers that likewise focus on long-term expectations and whether these are anchored include Beechey, Johannsen and Levin (2011), Garcia and Werner (2018), Eusepi et al. (2019), Galati, Moessner and van Rooij (2020), Yetman (2020), and Binder, Janson and Verbrugge (2019), while on the link between short and long run, see Andrade et al. (2016) and Mertens and Nason (2018).

others on disagreement across surveys, and yet others on difference in expected inflation across financial prices.[5] This paper instead focuses on disagreement between people and markets.[6]

The empirical model of household expectations combines elements from the literatures on imperfect information, overconfidence, learning, and sticky information, and builds more closely on Angeletos, Huo and Sastry (2020).[7] The model of financial markets builds more closely on Albagli, Hellwig and Tsyvinski (2013).[8] While these papers use the model to isolate a theoretical channel or to make predictions, here the model is a generalization that can be used as a measurement tool. As a result, the model is flexible enough to fit the expectations data exactly so that it can be used as a filter to detect the underlying fundamentals. Finally, the general-equilibrium model of inflation and monetary policy is a classic new Keynesian model, but set in continuous time building more closely on Reis (2019).[9] Finally, a different literature uses information on the yield curve and on inflation expectations from markets and surveys, to statistically decompose the discrepancy into different risk premia and nominal and real components.[10] This paper instead proposes a decomposition based on disagreement, the model has no parameters to estimate, and monetary policy affects and is affected by each component.[11]

2 A new variable of interest: the discrepancy

Let $\pi_{t,T}$ denote the change in the log of the price level between dates t and T. The discrepancy ϕ_t is defined as:

$$\phi_t = \mathbb{E}_t^*(\pi_{t,T}) - \mathbb{E}_t^p(\pi_{t,T}). \tag{1}$$

[5] For two examples of each, see Mankiw, Reis and Wolfers (2004) and Andrade and Le Bihan (2013), Carroll (2003) and Coibion and Gorodnichenko (2012), and Fleckenstein, Longstaff and Lustig (2014) and Andreasen, Christensen and Riddell (2017), respectively.

[6] See also Ang, Bekaert and Wei (2007) and Faust and Wright (2013) but from the perspective of a horse-race between the two, as opposed to explaining their difference.

[7] See, respectively, Woodford (2003*a*) and Angeletos and Lian (2016), Bordalo et al. (2020) and Guo and Wachter (2019), Malmendier and Nagel (2015) and Eusepi and Preston (2018), and Mankiw and Reis (2002) and Coibion and Gorodnichenko (2015).

[8] It is a version of the model in Grossman and Stiglitz (1980) surveyed in Vives (2008) or Veldkamp (2011); see Bassetto and Galli (2019) for a different application.

[9] The classic analysis is in Clarida, Galí and Gertler (2000) and Woodford (2003*b*).

[10] For instance, see Chernov and Mueller (2012), Haubrich, Pennacchi and Ritchken (2012), or Abrahams et al. (2016).

[11] Closer to this paper by also emphasizing disagreement is Cao et al. (2020).

The first term, $\mathbb{E}_t^*(.)$, is the expectation implicit in asset prices at date t, sometimes also referred to as: the risk-adjusted expectation, the expectation under the risk-neutral measure, or the break-even inflation. The second term, $\mathbb{E}_t^p(.)$, is the subjective expectation by the public as reflected in a measure of the central tendency of answers to surveys.

Measuring the discrepancy requires defining the country to which it refers to, the frequency of t, the horizon T, the asset price to extract the market measure, and the survey to measure the people's beliefs. Since the focus on this paper is on long-term expectation, T should be at least 5 years out, and given the focus on business-cycle movements, the frequency should be at least quarterly.

This still leaves many possible measures. For its baseline, the paper uses data for the United States, at a monthly frequency, and for a 5-year horizon. Data for market prices comes from inflation swap contracts, which are available daily since 2004. The swap market is quite liquid, and is heavily used by pension funds to hedge long-run inflation risk. Data for expectations by the public is from the Michigan survey of households, starting in 1978, taken as the median over the around 500 responses every month that answer: "By about what percent per year do you expect prices to go up/down on the average, during the next 5 to 10 years?".

2.1 Fact 1: the discrepancy has significant business-cycle fluctuations

The baseline monthly series, starting in 2010:1 and ending in 2019:4, demeaned, is plotted in figure 1. It moves around significantly across months with a standard deviation of 0.50%. For comparison, the standard deviation of actual inflation during that period is 0.57%, and these are long-horizon expectations.

Figure 2 confirms this by plotting in the top panel the spectral density of the discrepancy, as well as inflation's. In the grey box are the usual business cycle frequencies of 6-32 quarters. The discrepancy series has a significant amount of power in this frequency: 44% of its variation is accounted by the business cycle. In this, it resembles actual inflation, where 49% of the variation occurs at business cycle frequencies.

The middle panel investigates different series used to construct the discrepancy, plotting these series and their correlation with the baseline. Starting with the horizon, inflation swaps are also available for a 10-year horizon, so the alternative uses that for the market measure. A different market price comes from the difference between the yield in CPI inflation-indexed bonds (TIPS) and the yield in Treasury bonds.[12] A third alternative

[12] These are available for 5 or 10 year horizons, which are also the more liquid maturities, starting in 1999.

Figure 1: The discrepancy (market-people) over time

comes from the choice of series to measure the people's expectations. In the baseline, I took the median across respondents, whereas now I use the mean. Still focusing on the people's expectation, the fourth series uses data from the Survey of Professional Forecasters, which asks about CPI inflation over the next 5 and over the next 10 years. The survey is quarterly, but a monthly measure of its central tendency was calculated by the FRB Philadelphia.[13]

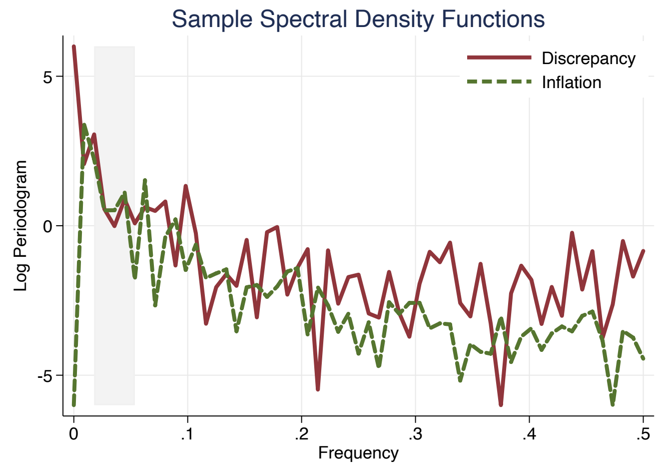
The figure shows that all of these series are highly correlated with our baseline, and at a business-cycle frequency the coherence between the two series is very high. Where they differ significantly is in their level and in their high-frequency variation. Inferences on whether market expectations are on average above or below those of people are suspect, and monthly (or even daily) monitoring of the discrepancy is unreliable, insofar as different measures can give very different answers. When it comes to the business-cycle movements that this paper focuses on though, the choice of series seems less important.

A main caveat of this measure is that for a significant part of the sample inflation was near zero, yet TIPS are not indexed to negative inflation, so that changes in this measure reflect changes in the option value embedded in this payoff, rather than changes in expected inflation.

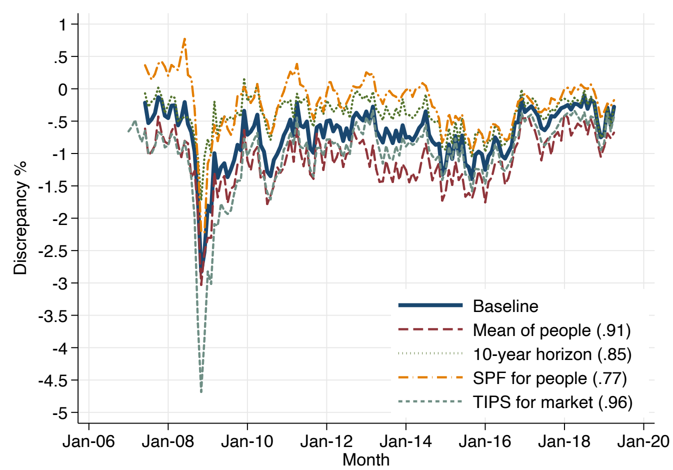
[13]This series includes business people and market participants; section 5 explores that distinction.

Figure 2: The business-cycle variation of the discrepancy

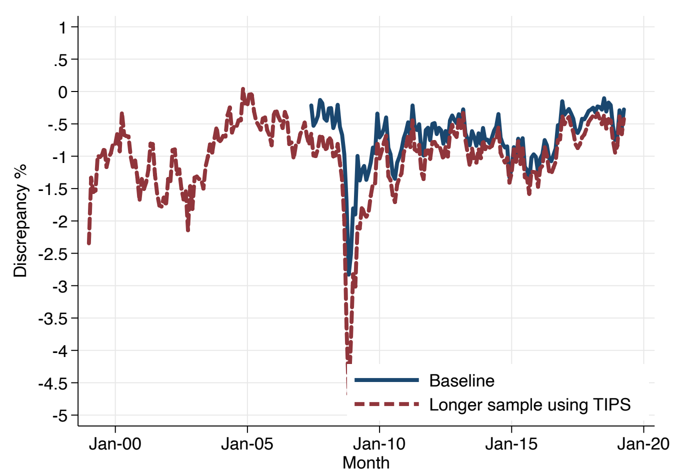
(a) Spectral density excluding frequency zero



(b) Alternative measures



(c) Longer sample



The previous panel plotted the data since August of 2007, when all the series are available. There is a clear large outlier at the end of 2008. The peak of the financial crisis significantly disrupted markets, and this especially affected the price of inflation-indexed bonds. The liquidity problems in that market, as well as the Fed's differential response in the nominal bonds versus the indexed-bonds market, may distort the measures of expected inflation. Therefore, in the baseline series, I start the sample in 2010, when these markets seemed to be operating normally again. To ensure 2008 was indeed exceptional, and to try to maximize the length of the sample, the bottom panel plots a longer alternative series, that uses TIPS for the market measures, since the data is available since 1999. The movements in the discrepancy between 2008 and 2009 are indeed exceptional. More interestingly, the business cycle movements in the discrepancy were also visible before 2008. The main difference is that the series further back in the past was slightly more volatile, as so was inflation.

More generally, the fluctuations in figure 1 are persistent, not just the results of high-frequency movements. Market liquidity frictions are important for daily movements, but they do not seem to interfere with the monthly or quarterly movements in the series. The correlation across series from different markets are very high at business-cycle frequencies even if they are much lower at a daily frequency.

2.2 Fact 2: the discrepancy is systematically related to monetary policy

The second column of table 1 shows estimates of a regression of the discrepancy on a summary indicator of the stance of monetary policy: the 2-year yield. Higher interest rates are strongly positively correlated with the discrepancy. Since this is usually negative, this means that tight policy lowers the difference between people and markets. This is consistent with market expected inflation rising by more than the people's expectation. The regression controls for inflation and its squared change.[14]

The third column instead regresses the discrepancy on a shock to monetary policy, constructed from movements in the Federal Funds forward market, from Nakamura and Steinsson (2018). Again, tighter unexpected changes in policy reduces the discrepancy between the markets and the people.

Figure 1 extends the analysis across countries. For the Eurozone, there are no indexed

[14] There may be some correlated noise in the contemporaneous 2-year nominal rate and the 5-year market measure of expected inflation. Using instead the lagged interest rate in the regression gives a very similar estimate of 0.139 with a 0.028 standard error.

Table 1: The proximate determinants of the discrepancy

	Determinants	Policy shocks
	(1)	(2)
2-year yield	0.149***	
	(0.0273)	
Inflation	0.177***	
	(0.0233)	
Squared change inflation	-0.200	
	(0.159)	
Monetary shocks		6.717
		(3.884)
Observations	111	43
R-squared	0.512	0.068

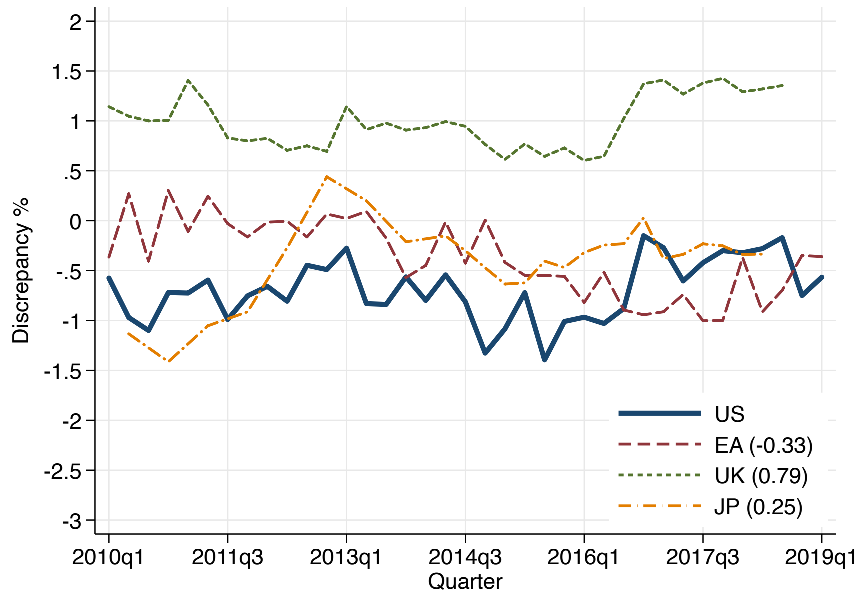
Discrepancy = markets - people; positive shock is tighter policy.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

bonds at the eurozone level but there are liquid inflation swaps, both at the 5 and the 10 year horizons, with daily data since 2004. For the people's expectations, an imperfect measure comes from the ECB's survey of professional forecasters, which asks respondents every quarter since 2004 for their inflation expectations over the long run. The figure shows the discrepancy using the 5-year swaps and the median of respondents to be closer to the US numbers. The United Kingdom has active and liquid markets for both inflation swaps and indexed government bonds at both the 5 and the 10 year horizon, the former since 2007 and the latter since 1997. For the people's expectations, I use the median of the Bank of England's 3-year ahead Survey of Economic Forecasters, which is quarterly and collected since 1998. Finally, for Japan, there are market prices for 5-year and 10-year indexed bonds and swap contracts since 2009. Subjective expectations come from Consensus economics, which has quarterly data on 10-year and 5-year ahead inflation expectations since 1989.[15]

All series are in figure 3 together with a quarterly version of the baseline series that is not de-meaned. While it is well-known that actual inflation is highly correlated across these countries, the correlation of the discrepancies is negative (for the EA) or low (for

[15] A caveat of all of these subjective measures is that they mix market participants and industry observers, unlike our baseline series from the Michigan survey, which surveys households directly. Section 5 discusses this further.

Figure 3: The discrepancy (markets - people) across countries



Japan), and their average level has different signs for different countries. Policies across these four regions were different during this time. Movements in the Japanese series match some of the changes in the policy regime there, from the introduction of qualitative and quantitative easing to yield curve control. In the UK, the end of sample shows a rising discrepancy following the Brexit referendum. In short, different policy regimes across these countries are correlated with different behaviors of the discrepancy.

2.3 Fact 3: a decomposition points to disagreement

Define $\mathbb{E}_t^b(.)$ and $\mathbb{E}_t^m(.)$ as the subjective expectations of bond-traders, respectively of the average and of the marginal trader. These can be different from each other, and from the subjective expectations of the households answering surveys $\mathbb{E}_t^p(.)$, because each individual forms her own expectations using her information and her beliefs. One can then decompose the discrepancy into three terms:

$$\phi_t = \underbrace{\mathbb{E}_t^b(\pi_{t,T}) - \mathbb{E}_t^p(\pi_{t,T})}_{\text{disagreement across}} + \underbrace{\mathbb{E}_t^m(\pi_{t,T}) - \mathbb{E}_t^b(\pi_{t,T})}_{\text{disagreement within}} + \underbrace{\mathbb{E}_t^*(\pi_{t,T}) - \mathbb{E}_t^m(\pi_{t,T})}_{\text{risk compensation}}. \tag{2}$$

The first term captures disagreement across types of agents, namely market traders and the public. We might expect the first group to be better informed about inflation, since it is part of their job, although they may also be more susceptible to fads, conformity biases, and short-term thinking. The second term is also about disagreement, but now between the marginal trader in the market, whose views prices reflect, and the average trader. Heterogeneity of views in markets is what causes trade in the first place. At the same time, imperfections in financial markets arising from liquidity shocks, changes in risk-taking capacity across agents, and shifts in heterogeneous levels of confidence together with short sales constraints, would show up here. Finally, the third term is the pure compensation for inflation risk that this individual marginal trader will require if she is risk averse.

There are empirical proxies for the first and the third term. Starting with the former, the FRB New York has since 2010 surveyed about 50 financial market dealers eight times per year on their expected inflation over the next 5 years.[16] These are dealers in Treasury markets trading inflation risk, therefore matching closely the $\mathbb{E}_t^b(.)$ concept.[17] Turning to the latter, the arguments in Martin and Wagner (2019) for equity risk applied to the market for inflation risk, suggest that the risk-neutral variance of inflation measures compensation for inflation risk. This can be measured using options for inflation as in Hilscher, Raviv and Reis (2014).

Figure 4 plots the discrepancy and these two measures, while table 2 regresses the discrepancy onto the two components. While the disagreement between the beliefs of traders and households can account for a significant part of the time-series variation in the discrepancy, risk compensation instead quantitatively accounts for very little.

This regression treats the within disagreement term as a residual. Directly observing who is the marginal trader is a hopeless empirical task. Moreover, it is hard to believe that this omitted term is orthogonal to the other two. More generally, the three terms in the decomposition are correlated with each other and likely depend on the same factors. While direct risk compensation accounted for little of the discrepancy, the factors that drive it

[16]The survey also includes quartiles of the distribution, and for 2010-12 only asked about an expectation of inflation starting in 5 years' time for the next 5 years. However, these extra series add little additional information.

[17]An alternative is to use the sub-sample of respondents that work for financial firms on the Survey of Professional Forecasters: the correlation between that disagreement-across series and the one I use is 0.93. An inferior alternative is the Blue Chip survey, which surveys mostly financial market participants, but across many different types of financial markets, many of them far from markets where inflation risk is actively traded.

Figure 4: The decomposition of the US discrepancy

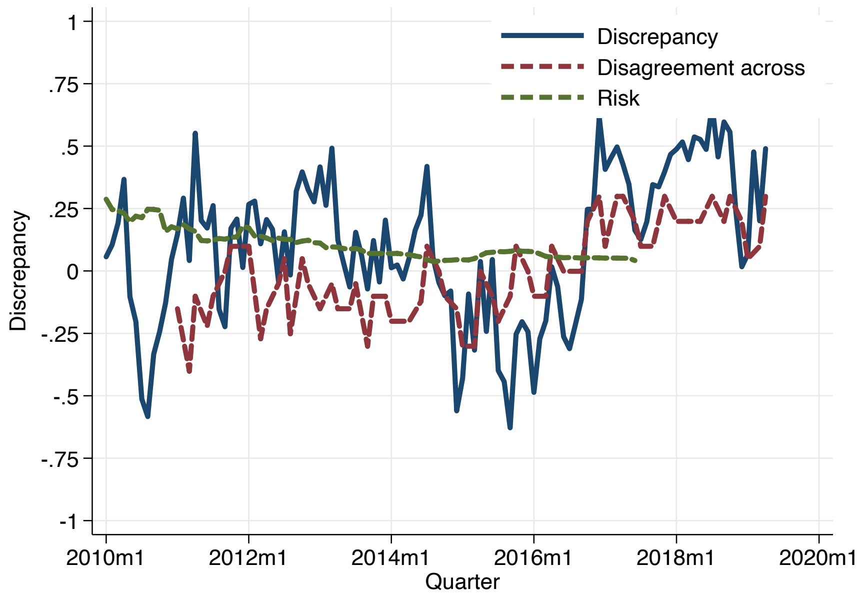


Table 2: The discrepancy and its components

Terms	Coefficients
Disagreement across	0.746***
	(0.205)
Risk compensation	1.904**
	(0.814)
Observations	52
R-squared	0.270

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

(or that drive risk understood more broadly) may be the determinants of disagreement. To make progress, one needs a model of financial markets and heterogeneous beliefs that can provide estimates of the within-disagreement term as a function of observables, and relate the three terms to the shocks that drive them. The rest of the paper does this.

2.4 Inflation risk premia

Before proceeding, it is worth clarifying how risk premia fit into the discussion. Let $\pi^e_{t,T}$ denote the *rational expectation* of inflation. Then, the inflation risk premium is defined as:

$$\text{IRP}_t \equiv \mathbb{E}^*_t(\pi_{t,T}) - \pi^e_{t,T} = \underbrace{\phi_t}_{\text{discrepancy}} + \underbrace{\mathbb{E}^p_t(\pi_{t,T}) - \pi^e_{t,T}}_{\text{irrationality}}. \tag{3}$$

As the equality shows, this is not the same as the discrepancy. Only if there are homogenous rational expectations would they be the same, in which case the inflation risk premium would be equal to the risk compensation in the decomposition in figure 4. Not only has the hypothesis of rational expectation been strongly rejected many times using our Michigan survey data (and many other datasets), but the pure risk compensation accounted for a small part of the variation in the discrepancy.[18]

Still, taking ϕ_t as a proxy for an inflation risk premium, the decomposition above suggests that inflation risk is quantitatively driven more by disagreement between different traders in financial markets. Departing from the benchmark of Milgrom and Stokey (1982), a voluminous literature has considered models with imperfect information, behavioral biases, liquidity shocks, short sales constraints, among many others that give rise to inflation risk due to this disagreement. These are consistent with figure 4, as they are models of disagreement within and between rather than pure individual-level risk compensations. Section 5 will derive estimates of $\pi^e_{t,T}$ with which one could test those models.

3 A parsimonious model of the people's expectations

The anchor of the people's model of expectation is the rational expectations fundamental $\pi^e_{t,T}$. The only assumption made on it is that it is unbiased and leads to serially uncorrelated forecast errors: $\mathbb{E}_t(\pi_{t,T}) = \pi^e_{t,T}$, and $\mathbb{E}_t(\pi^e_{t,T}(\pi_{t,T} - \pi^e_{t,T})) = 0$ if $\mathbb{E}_t(.)$ is a statistically

[18] See also Gürkaynak, Sack and Wright (2010).

optimal expectation operator. People do not know what this is. Indexing a household by h, and dropping the time subscript (t, T) that would otherwise appear everywhere, its individual expectation is denoted by v^h.

3.1 Four behavioral features of expectations

The first property of household expectations is *incomplete information*. Starting from some common prior π^*, each household receives an idiosyncratic noisy signal drawn from a distribution centered at π^e but with variance σ^2. This signal induces a dispersion of expectations across people as a result of different signals being drawn. At the same time, because households know their signals are noisy, this imperfect information leads average inflation forecasts across households to under-react to news, as a large literature has found (e.g. Coibion and Gorodnichenko, 2015). I impose a simplifying assumption: that the distribution of signals is normally distributed.[19]

The second property of expectations is *over-confidence*. Households behave as if their signals are more precise than what they really are. Therefore, if the response of their individual expectation to a signal is given by δ, this can be high even when $1/\sigma^2$ is low, and could even be above 1, something a rational imperfectly informed agent would never choose to do. This matches the also extensive literature that individuals, in the cross-section, over-react to signals (e.g. Bordalo et al., 2020). The simplifying assumption in modeling this is to assume that the relation is (at least approximately) linear.[20]

Third, households *learn from experience*. Namely, they suffer from a bias z_c in their beliefs, that arises from past experiences that left a scar. Empirically, a growing literature has found that experiences of inflation, especially at younger ages, account for a significant share of the disagreement observed across age cohorts (e.g. Malmendier and Nagel, 2015). The simplification here is to assume that this bias is linear in the age of the cohort, so if $c = 0, 1, 2...$ denotes the cohort, then $z_c = c\pi^z$ for some constant π^z. This matches the fact that in the US, inflation has trended down since the 1980s, so that as one moves to older cohorts, so c rises, the bias is higher.[21]

[19] This would be optimal if households had a quadratic objective function, the prior was normal, and they suffered from rational inattention (Sims, 2003).

[20] If the overconfidence shows up as the perceived variance of signals being smaller than the actual variance, then with a normal signal, the linearity follows from the properties of the conditional normal distribution. Overconfidence can follow as well from believing the long-run mean of the series if affected by the signal (Afrouzi et al., 2020)

[21] If inflation followed a random walk with white noise, that happened to have mostly negative permanent shocks in the last 40 years, then a least squares learning formula that is infrequently updated would

Finally, cohorts update their expectation bias according to *sticky information*. Cohorts do not refer to age, but to time since the last bias update. Every period, a small share of households of any given cohort updates its information and eliminates the bias it had from the past. This leads to a slow dissemination of information, and to disagreement that evolves endogenously with the shocks to inflation, as observed in the data (e.g. Mankiw, Reis and Wolfers, 2004). I make the usual simplification in this literature that this process evolves according to a memoryless Poisson process, so that at any date in time there is a fraction $\lambda(1-\lambda)^c$ of people that have a bias according to cohort c.[22]

These four properties capture the main features of a significant share of the research on modeling expectations over the last two decades. Together with the simplifying assumptions introduced along the way, they give rise to the following parsimonious empirical model of expectations, where the time subscript t is re-introduced (the horizon is always T):

$$v_t^h = c_t \pi_t^z + \pi_t^* + \theta_t(e_t^h + \pi_t^e - \pi_t^*) \tag{4}$$

$$e_t^h | \pi_t^e \sim N(0, \sigma_t^2) \text{ and } c_t \sim Exp(\lambda_t) \tag{5}$$

The model has four parameters capturing the strength of the four behavioral mechanisms described above: σ_t^2 on how disperse and imperfect is information, θ_t on how over-confident people are, π_t^z on how large are the scars of past high inflation, and λ_t on the stickiness of information on updating biases. Conditional on the two unobservables, the prior π_t^* and the actual fundamental π_t^e, the model predicts that the observable individual expectation v_t^h has a distribution $F_t(.)$ that is an exponentially-modified Gaussian. Its first three moments are non-zero. With data on the average, the standard deviation (or interquartile range), and the skewness of inflation expectations in the household survey, the model identifies three parameters: $\sigma_t^2, \theta_t, \lambda_t/\pi_t^z$.

3.2 Application to the Michigan survey data

Figure 5 fits this model to the data on long-term inflation expectations in the Michigan survey since 1990. The top panel shows the shape of $F_t(.)$ when using the average over

generate this process endogenously.

[22]The cohorts and their biases can be given a broader interpretation. For instance, perhaps they are driven by differences in consumer baskets, education, or exposure to salient prices like gas prices. The key assumption here is that the cross-sectional distribution across these types is exponential, as if membership in each type was updated following a Poisson process.

time of the mean, standard deviation, and skewness in the data. The distribution looks like a normal distribution, but it has a fatter tail on the right reflecting the presence of the upward bias due to the scars of the 1980s when US inflation was higher.

The middle panel shows the three moments in the data over time. As has been noted before, the average long-run inflation expectation fell throughout the 1990s, before stabilizing around 3%. Some interpret the fall as the success of expectations anchoring, while others point to it stabilizing a full percentage point above the target of the Federal Reserve as a failure. Less appreciated is that the standard deviation of expectations was roughly constant throughout, but then started falling around 2014. At the same time, the skewness also slightly fell, remaining positive but lower than it had before.

The bottom panel shows the implied σ_t^2 and λ_t / π_t^z to match these moments using the model. According to the model, after a spike in 2008-10, there is a visible trend downwards in the dispersion of information that, by the end of the decade, is persistently at a very low level. At the same time, the bias arising from cohorts infrequently updating has fallen significantly from 2014 onwards, as either the size of that bias π_t^z has fallen, or the frequency with which agents update and reduce it λ_t has risen.[23]

With these parameters varying over time, the model fits the first three moments of the expectations data exactly. At the same time, there are three over-identification tests of whether the model fits the overall data well.

The first is that the model imposes that the two series on the bottom panel of the figure have to be positive. In the expectations data, this turns out to be the case at every single date. This was not guaranteed, since if the skewness was much higher at any one date, the implied σ_t^2 at that date would have been negative. The fact that it never is so supports the setup of the model.[24]

Second, the model predicts that kurtosis (and higher-order moments) should be zero. In the data, on average that is approximately the case over the last thirty years. Focussing on the first three moments, as the model does, is a good statistical approximation.

Third, the model predicts that the following peculiar combination of data statistics should be equal to the average prior over time:

$$\mu_t \equiv Mean_t - StDev_t(0.5Skew_t)^{1/3} \quad \Rightarrow \quad \lim_{T\to\infty} \frac{\sum_t \mu_t}{T} = \pi^*. \tag{6}$$

[23] The model does not identify the remaining parameter, θ_t, independently of the fundamental π_t^e.

[24] As a validation, in the Survey of Professional Forecasters subsample of financial participants, this is not the case, partly because skewness is sometimes negative.

Even though the average of $Mean_t$ is 3.59%, the left-hand side of this expression is 2.26%. Post 2010, the average long-run inflation expectation in the Michigan survey is 3.11%, but according to the model π^* will have been only 1.92%. That is, without any free parameters, the model can make sense of the high survey expectations reported in the surveys as being consistent with a plausible value for the underlying long-run average of implied expected inflation. It is the experience scar that leads to the high reported survey answers.

To be clear, the fit of this parsimonious empirical model does not by itself provide a test of each of the four behavioral assumptions and associated simplifications. What it shows is that the model is a good measurement tool with which to filter the noisy expectations data and that is broadly consistent with the major insights from micro-founded models.

4 A parsimonious model of financial market expectations

Consider a continuum of traders indexed by i in the unit interval, whose subjective beliefs are drawn from a cross-sectional distribution $P_t(.)$, that is different from the $F_t(.)$ distribution of household beliefs.

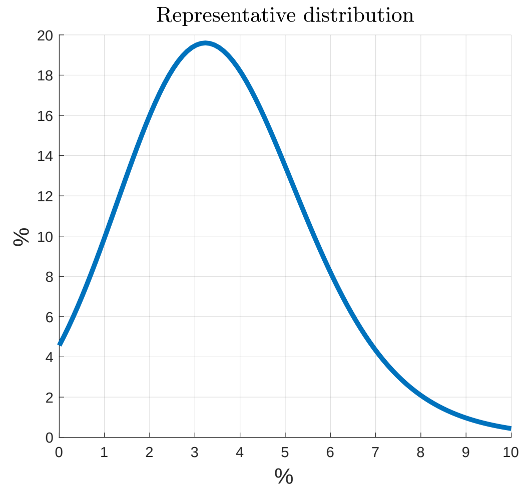
I assume that traders are drawn from the population of households, so they enter markets with a prior expectation of inflation that is a draw v^i from the posterior $F(.)$ distribution (again omitting time subscripts). Traders understand how this prior reflects the fundamental, that is they understand the model and its biases given by equations (4)-(5). Moreover they observe an extra piece of information: the market price q of a nominal bond that next period gives 1 nominal unit next period. Prices are not fully revealing but they reveal some information on what the fundamental is, which is captured by the distribution $g(q|\pi^e)$. Traders are Bayesian, so their posterior for inflation is:

$$p(\pi^e|v^i, q) = \frac{g(q|\pi^e)f(\pi^e|v^i)}{\int g(q|\pi^e)f(\pi^e|v^i)d\pi^e}. \tag{7}$$

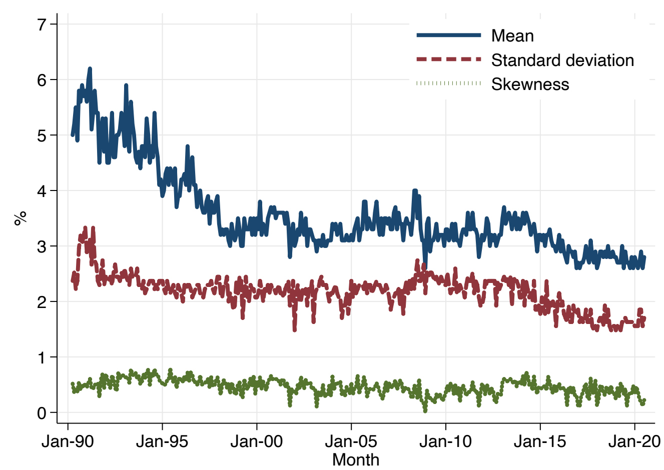
A plausible alternative to this specification would be to assume that traders have also differ in their non-price subjective information, so that their prior v^i would be drawn from a different distribution than the one that fits the Michigan data. Replacing all the references to the $F(.)$ distribution in this section by that traders' distribution, the model would be unchanged. However, estimating this new distribution is not feasible. The FRBNY survey of dealers surveys about 50 people; far too few to estimate the skewness

Figure 5: The model of households expectations and the data

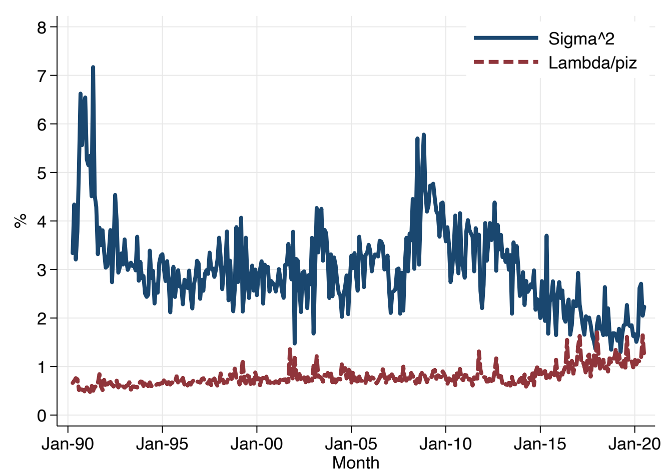
(a) Representative distribution $F(.)$



(b) Three first moments of Michigan household survey



(c) Parameters over time



accurately, or even for the estimates of the second moments to be reliable.[25]

Another alternative would be to assume that traders instead have a prior distribution for π^e that matches the rational expectations fundamental, centered at π^*. They would then each get matched with one household, and learn their expectation v^i. But, they would treat it as a signal as opposed to taking it as their prior, as I assumed above. The appendix writes this alternative model. Its main properties are the same as the model I use, as are tits resulting estimates on US data.

4.1 Individual behavior

The goal of each trader is to maximize the expected discounted profits by choosing bond-holdings b^i:

$$\max \int \left[m(\pi)e^{-\pi} - q\right] b^i p(\pi^e|v^i, q)d\pi^e, \tag{8}$$

where $m(\pi)$ is the stochastic discount factor, that depends on inflation insofar as there is aversion to inflation risk. I impose two simplifying assumptions on this problem. First, each trader has some positive wealth w^i, and cannot borrow or short the bond, so that $b_i \in [0, w_i]$. Second, the stochastic discount factor is common across all traders, given to them by the head household to whom they return their profits at the end of the trade. Since each of them is infinitesimal, then the payoff from investing in the bond is: $y(\pi^e) \equiv \mathbb{E}\left[m(\pi)e^{-\pi}|\pi^e\right]$ which does not depend on each individual trader. While there is risk aversion captured in the curvature of $y(\pi^e)$, each trader individually behaves as if she was risk neutral after adjusting for risk compensation.

These two assumptions, together with the fact that the exponentially-modified Gaussian distribution $F(.)$ has a monotone likelihood ratio, imply a simple solution to the problem. As long as the posterior after observing the price is not degenerate, then traders that expect high inflation at the start, $v^i > v^*$ do not want to hold the nominal bond $b^i = 0$; traders that expect low inflation $v^i < v^*$ invest all their wealth in the bond $b^i = w^i$, and the marginal trader who is just indifferent defines the threshold v^* as:

$$\int y(\pi^e)p(\pi^e|v^*, q)d\pi^e = q. \tag{9}$$

[25] The Survey of Professional Forecasters mixes financial market participants with other businesses; separating between the two leaves again too few financiers to reliably estimate second and third moments. Finally, the Blue Chip survey median of long-run expectations does not correlate very well with that from the dealers', perhaps because many of its financial-market participants are quite removed from trading inflation risk, unlike what happens in the FRBNY survey.

4.2 Market clearing and noise

Given a total supply of bonds B, the standard market clearing condition is $\int_0^1 b^i di = B$. Given the threshold strategy for investment just derived, then letting w be average wealth across agents, the market clearing condition becomes the simple condition:

$$F(v^*|\pi^e) = B/w \equiv \omega. \tag{10}$$

If wealth and the supply of bonds were known, then the market price would perfectly reveal to every trader what the actual fundamental expected inflation π^e is.[26] Their posteriors $p(\pi^e|v^i, q)$ would all be identical and degenerate (Grossman and Stiglitz, 1980), and they would be indifferent between trading the bond or not (Milgrom and Stokey, 1982). A long literature has broken this homogeneity of beliefs and resulting efficiency of markets by arguing for different behavioral assumptions on how traders fail to solve this inference properly, different market structure assumptions that lead to noise traders and time-varying frictions that map into shocks to w, and shocks to the supply of bonds or to their liquidity that map into B.[27] A parsimonious model of these imperfections is to assume that ω is a random variable that the traders cannot observe. It is then a source of noise contaminating asset prices.[28]

The range of this variable ω is the unit interval. I make the simplifying assumption that it follows a symmetric Beta distribution with parameter β. When β approaches 1, the distribution approaches the uniform, while when it grows to infinity, it becomes concentrated in a point mass at $1/2$. For any finite value, it implies that market prices will reflect both the fundamental as well as this noise, matching the high volatility observed in the prices of inflation swaps.

[26] Note that only average wealth matters, and not its distribution, because the signal drawn is independent of wealth, and all the traders with a signal below V^* choose to put all of their investment into the bond.

[27] A few examples of this literature that dates back to Harrison and Kreps (1979) are Harris and Raviv (1993), Scheinkman and Xiong (2003), Shefrin (2008), Dumas, Kurshev and Uppal (2009), Xiong and Yan (2010), and Simsek (2013).

[28] It is easy to relax the assumption of no short-sale constraints. Imagine there were such a constraint of $\underline{b}_i$ per trader, with an average $\underline{b}$ across traders. The marginal trader condition in equation (9) would be unchanged, as would the market clearing condition in equation (10), but now $\omega = (B - \underline{b})/(w - \underline{b})$. Tightening and loosening of the short-sales constraint would be a further source of financial shocks affecting market prices.

4.3 Equilibrium

A perfect Bayesian equilibrium in this financial market is a price function $q(\pi^e, \omega)$ such that the beliefs of each trader follow Bayes rule in equation (7), their investment choices satisfy the optimal threshold strategy in equation (9), and the bond market clears in equation (10). Following Albagli, Hellwig and Tsyvinski (2013) though, inspecting the three equations that determine equilibrium, it is clear that the threshold v^* is a sufficient statistic for the pair (π^e, ω) to determine prices. Letting $q = Q(v^*)$, note that only v^* appears in the belief function in equation (7) and in the investment threshold function in equation (9). As for the market clearing equation (10), one can invert the $F(.)$ distribution function to obtain: $v^* = \pi^e + F^{-1}(\omega)$, using the fact that the signals in the survey distribution are centered around the fundamental π^e. Given the Beta distribution for ω, this equation produces a distribution function for the threshold that is consistent with markets clearing: $G(v^*|\pi^e)$.

The solution of the model is therefore the solution of the equation:[29]

$$Q(v^*) = \frac{\int y(\pi^e) g(v^* - \pi^e) f(v^* - \pi^e) d\pi^e}{\int g(v^* - \pi^e) f(v^* - \pi^e) d\pi^e}. \tag{11}$$

Given a solution for $Q(.)$, recovering the solution for the original model follows from: $q(\pi^e, \omega) = Q(\pi^e + F^{-1}(\omega))$.

The model and its solution have a few properties that make it useful to understand market-implied inflation expectations as they relate to survey expectations. First, an increase in the prior mean π^* shifts the predicted asset price q one-to-one. Therefore the prior, which is a free parameter of the model, anchors the results in a transparent way. The other free parameter is β, which determines how informative asset prices are, again transparently accommodating the views of different researchers.

Second, the asset price is monotonic in both the underlying fundamental π^e and noise ω. Therefore, higher expected inflation according to markets may either reflect signal or noise. It is easy to show that when ω approaches its lower (upper) limit of 0 (1), then the threshold v^* approaches minus (plus) infinity, and likewise for the price. Therefore, for any observed market price, the model can always make sense of it in terms of a realization of noise, even if very unlikely. Like the model of people's expectations, this model of market expectations can always fit the data exactly.

[29]Note that there was a slight abuse of notation in using $g(q|\pi^e)$ in equation (7) and then replacing it with $g(v^*|\pi^e)$.

Third, the model has a natural mapping to the empirical objects in section 2. For starters, the market expectation of inflation is $\mathbb{E}^*(\pi^e) = 1/q(\pi^e, \sigma) - (1+r)$ where r is a measure of the safe real rate. Next, replacing the payoff function with a linear approximation: $y(\pi^e) = \bar{y} + \pi^e$, then $\bar{y}$ is the *risk compensation* under homogeneous rational expectations. Adjusted for risk, the asset price is the expectation of inflation according to the marginal trader, or $\mathbb{E}^m(.) = \mathbb{E}(.|v^*, q)$. In turn, letting the median of the $G(v^*|\pi^e)$ distribution be v^{med}, then the median in a survey of traders would be:

$$\mathbb{E}^b(\pi) = \frac{\int \pi^e g(v^* - \pi^e) f(v^{med} - \pi^e) d\pi^e}{\int g(v^* - \pi^e) f(v^{med} - \pi^e) d\pi^e}. \tag{12}$$

The *disagreement within* term is then the disagreement between the trader with prior expectation (before observing market prices) v^*, and the trader with prior expectation v^{med}. Finally, the *disagreement across* is due to traders observing prices, while households do not. So, traders' expectations follow the distribution distribution $P(.)$, while households' expectations are drawn from $F(.)$. Therefore, the model fits perfectly into the decomposition of the discrepancy in section 2.3. Because the expectation of the marginal trader is unobservable with data alone, and v^* is an equilibrium object, the model is able to complete the decomposition.

Combining all of these properties, adding financial markets allows the parsimonious empirical model to now also fit market-based expectations of inflation, and to explain the discrepancy in terms of its decomposition into underlying terms. The model makes predictions for the three moments of the survey distribution of households $F_t(.)$, the median expectation of traders, and the market-implied expectation in the price q_t. They identify the five unknowns, three from the behavioral model of expectations $\theta_t, \sigma_t^2, \lambda_t/\pi_t^z$, one from the model of financial markets ω_t, and finally the fundamental expected inflation anchor π_t^e. The model therefore provides a measurement tool—a computational filter—through which to measure fundamental expected inflation, and thus to assess whether long-run inflation expectations are truly anchored.

5 Measuring long-run inflation expectations

There are many expected inflations in the model. Markets, people, and traders all deviate from the fundamental expected inflation π^e, but they are all anchored by it. Moreover, in a general-equilibrium model (like the next section will illustrate) this fundamental is the equilibrium object that may be anchored or not depending on policy. Therefore, answering the question in the introduction—Are inflation expectations anchored?—boils down to estimating π^e. In turn, estimating the marginal trader v^*, decomposes the discrepancy into disagreement within and across.

5.1 The model's mechanics

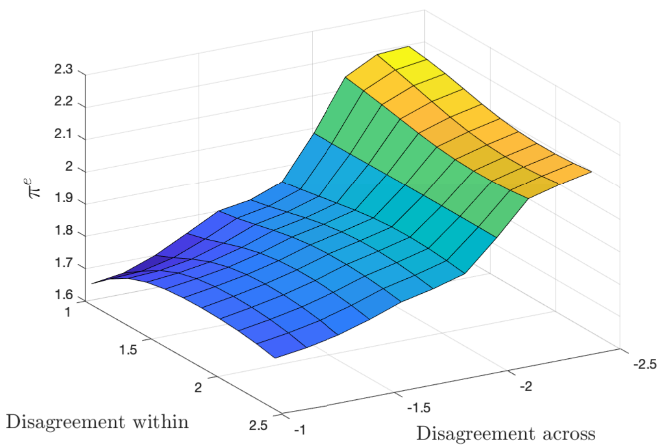
The model has only two parameters at each date: π_t^*, anchoring the fundamental, and β on the volatility of the noise. If the prior for inflation expectations is taken as constant, then it plays no role on the dynamics of the discrepancy and of expected inflation, since it moves all expectations precisely by the same amount. I take this to be the baseline case, and choose $\pi_t^* = 2\%$. This was the announced target of the Federal Reserve, and assuming a fixed prior on it, if anything biases the results towards finding anchored expectations at this target. The appendix reports an alternative where instead one year's prior is taken to be the previous year's estimate, that is: $\pi_t^* = \pi_{t-1}^e$. The model then becomes a dynamic filter updating its prior every period to be last year's posterior. The estimates turn out to be quite similar. As for β_t, I set it equal to 2, so the prior on market noise is reasonably flat. The appendix reports estimates with 1.1 or 4, which turn out to be quite similar.

Figure 6 shows how the model's predictions for expected inflation and the marginal trader vary with changes in the two sources of disagreement. The figures vary the market expected inflation and the survey expected inflation, while keeping the traders' average belief fixed. As disagreement across becomes more negative, that is, as people expect higher inflation while traders' expectation is unchanged, then the model predicts that fundamental expected inflation must be higher. In the other direction, if we observe traders' expectation falling relative to those of people, the model will signal a fall in fundamental expected inflation. Quantitatively, ceteris paribus, higher people's expectations have a stronger impact on estimates of the fundamental than higher market expectations. This is because the model attributes some of the disagreement within to being caused by the financial shock, and increasingly so when the disagreement within gets larger.

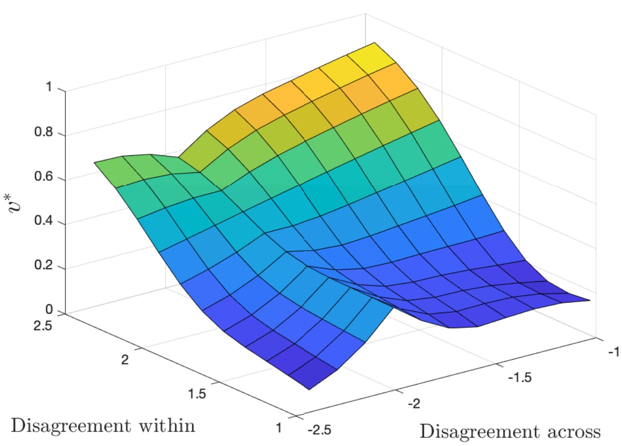
Further, as disagreement within becomes more negative, so the asset price falls further below the average trade's belief and the discrepancy rises, then the bottom panel shows that the marginal trader moves towards the left tail away from 0.5. The model partly interprets this as a result of noise, so the fall in expected inflation is muted but still positive. If instead disagreement within is positive, while disagreement across was nega-

Figure 6: The model at work

(a) The model's predicted expected inflation



(b) The model's predicted marginal trader



tive, that is if the average household expects higher inflation than the average trader, but the market-prices are also above those of traders, then the model interprets this partly as a result of noise pushing prices up, but still increases its estimate of expected inflation. When it comes to identifying the marginal trader, the disagreement within is particularly informative.

5.2 US estimates post-2011

The input data series are the three moments of the Michigan survey, the market prices from inflation swaps as in the baseline discrepancy series, the implied variance in options contracts to measure risk compensation, and the median from the FRBNY survey of dealers.[30] The latter series starts in 2011 and is not available every month, since the survey takes place only 8 times per year. Moreover, as figure 5 showed, the survey data is quite noisy month-to-month. Therefore, I average each of the data input within each of the nine years of the sample, and solve the model nine times, once per year.[31] The model matches each of these series exactly, and produces the underlying series for π^e and v^*.

The top panel of figure 7 shows the implied fundamental expected inflation. This is the unique solution to the model: the only series consistent with the the data from the people, the markets and the traders. It shows a clear downward trend from 2014 onwards. Starting from 2%, long-run expected inflation has fallen to around 1.8% by 2019.

The middle panel shows who is the marginal trader at each date in time. Strikingly, between 2013 and 2016, the marginal trader goes from being roughly the same as the average trader to being the trader on the 20th percentile left tail. To explain the fall in market-implied expectation when survey expectations moved little, the model points to the emergence of negative skewness among traders, with a significant share of them expecting very low inflation. Then, after 2016, it is skewness among the households that falls (recall the large rise in λ/π^z in figure 5), becoming less positive than before, solidifying the downward trend in expected inflation. First in markets, then in households, the skewness reveals a significant number of people expecting inflation to be significantly lower than the 2% inflation target.

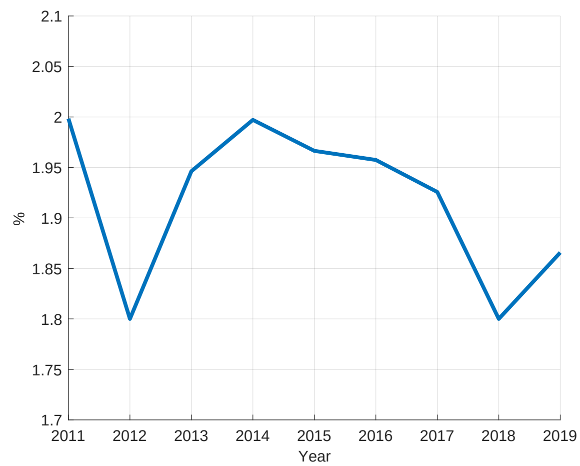
The bottom panel shows the decomposition between the two sources of disagree-

[30] See footnote 25 for why other surveys of professionals do not serve the purpose.

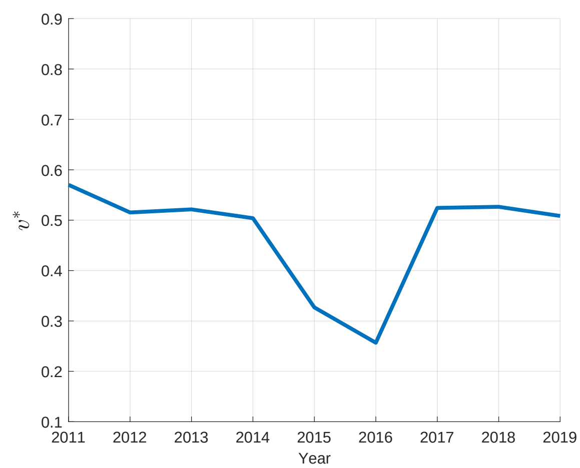
[31] The appendix reports the results of solving the model for every month that the data is available. The annual average of the estimates is similar to the estimates using annually averaged data. Moreover, the month-to-month fluctuations are erratic and driven by the jumps on the moments of the survey distribution.

Figure 7: Model estimates for baseline US sample

(a) Fundamental long-run inflation expectations



(b) The percentile of the marginal trader



(c) The decomposition of the discrepancy over time

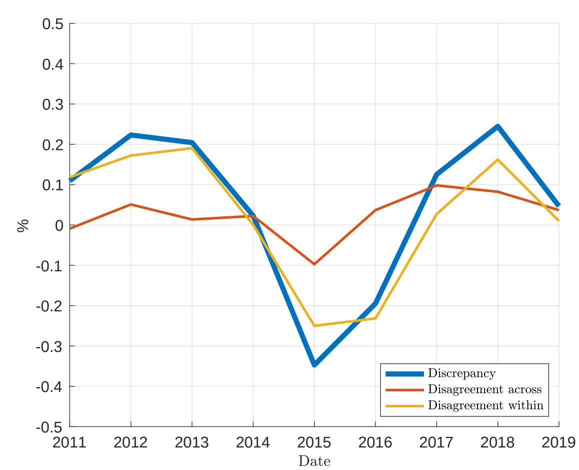
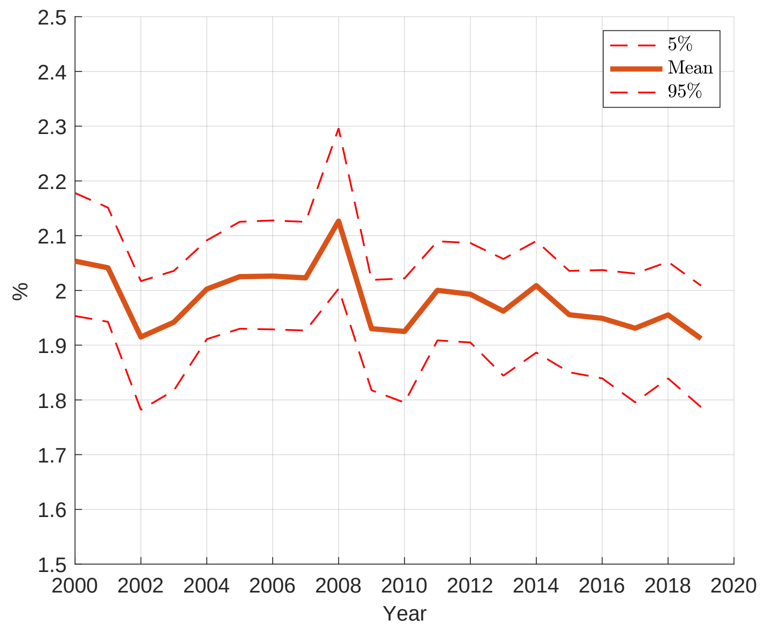


Figure 8: Estimates of expected long-run US inflation since 2000



ment, which confirms this interpretation. It is the disagreement within market traders that drove the discrepancy down in 2013-16, with the marginal trader moving further away from the median. This reverts from 2017 onwards, but by then there is a fall in expectation of the people, which keeps expected inflation down by closing some of the disagreement across, and reducing the discrepancy.

5.3 US estimates post-2000

Before 2011, there is no reliable data for the expectation of traders. There is however data from the survey of households, and from market prices using TIPS since the start of 2000. With this missing data, the model no longer pins down a unique fundamental expected inflation at every date. Instead, at each date, there is a pair (π_t^e, ω_t) that is consistent with the data. There is now an inference problem.

I solve it by taking a Bayesian approach: choosing priors for the two unknown series, and using the model to produce posteriors. Like the agents in the model, a Beta distribution for ω provides a prior for it. Also like the agents in the model, but removed of their behavioral biases, I choose a normal prior for π_t^e centered at π_t^* and with a standard deviation of 0.1. The model has one curious property, however. The parameter determining the sensitivity of household expectations to their signal is:

$$\theta_t = \frac{\mu_t - \pi_t^*}{\pi_t^e - \pi_t^*}. \tag{13}$$

Since $\theta_t > 0$, in the dates where the data implies that $\mu_t > \pi_t^*$, then the model immediately infers that $\pi_t^e > \pi_t^*$, and vice-versa when the sign switch. Therefore, the effective prior is a truncated normal, either to the left or to the right of π_t^*. As survey expectations move around so μ_t hovers around 2%, the prior moves considerably and this would move the estimates. To deal with this issue, instead of fixing $\pi_t^* = 2\%$ I set a hyper-prior for it, following a normal distribution with standard deviation 0.1%. This way, at every date, different values of π_t^* are considered while the data fixes μ_t. The prior switches, but the hyper-prior averages around these switches.[32]

Figure 8 shows the mean and two percentiles of the posterior distribution for expected inflation. There is still a fall after 2010, although the new prior imposing no change naturally makes it less dramatic. In the new sample, there is an interesting steady increase from 2002 to 2007, a period where arguably monetary policy was loose, at least relative to a Taylor rule. The model suggests that the common view that long-run inflation expectations have been essentially constant at 2% for the last 20 years is misleading. While the time-series changes have not been extreme, they are still clearly visible.[33]

5.4 The Eurozone

Looking for more data across regions, as opposed to over time, figure 9 shows the evolution of the discrepancy (top panel) and of the distribution of expectations in the Survey of Professional Forecasters for the Eurozone. The SPF is not the best sample for the model, since its forecasters are considerably more informed than the households but, at the same time, many of them are not traders that we could map into our model. There are no comparable data to estimate the model as was the case in the US. At the same time, as the top panel shows, the discrepancy has been particularly large and persistent in the EZ, with

[32]The appendix reports the robustness of the results to the value chosen for the standard deviation of the prior.

[33]The appendix reports the evolution of the percentile of the marginal trader: the marginal trader was to the right of the median in 2004-07, consistent with the discrepancy becoming smaller in absolute value then, but then it falls precipitously in 2008 matching the fall in market prices for expected inflation in the financial crisis.

market expectations stable around 1% since 2014, while the SPF expectations were stable at 1.9%.

Informed by the lessons from the model, the bottom panel provides hints at what is going on. The distribution of expectations shows a noticeable increase in variance between 2010 and 2014. More striking is the emergence of a clear negative skew. This suggests that the marginal trader in EZ inflation contracts was shifting noticeably to the left around the same time as it did in the US. However, unlike the US, in the following years, the mean of the distribution fell (as did its variance). This suggests that, in the first stage, disagreement within, followed in the second stage by disagreement across, kept market expectations low, and likely to a lower fundamental expected inflation in the EZ than in the US. Again the combination of movements in skewness and variance drive this conclusion.

6 Monetary policy and inflation with a discrepancy

This section embeds the partial-equilibrium model of financial markets into a general-equilibrium model of inflation determination. The goal is to endogenize the fundamental π_t^e, so as to understand the interaction between it, the discrepancy, and monetary policy. The model of monetary policy and inflation is the one at the core of most new Keynesian DSGE models, but is kept as simple as possible, leaving for future research the study of the many other possible interactions between these variables.

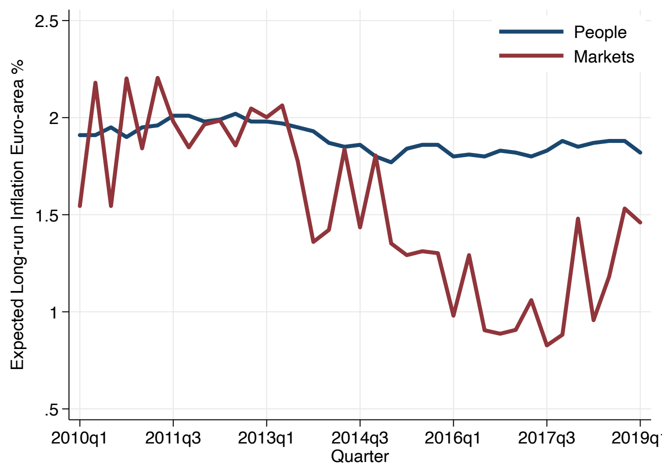
In terms of notation, time t is continuous, p_t is the price level, g_t is the expected growth rate of output, and Z_t is a vector of the shocks hitting the economy, which are two Ito processes. Inflation is given by:

$$\frac{dp_t}{p_t} = \pi_t^e dt + \alpha' dZ_t, \tag{14}$$

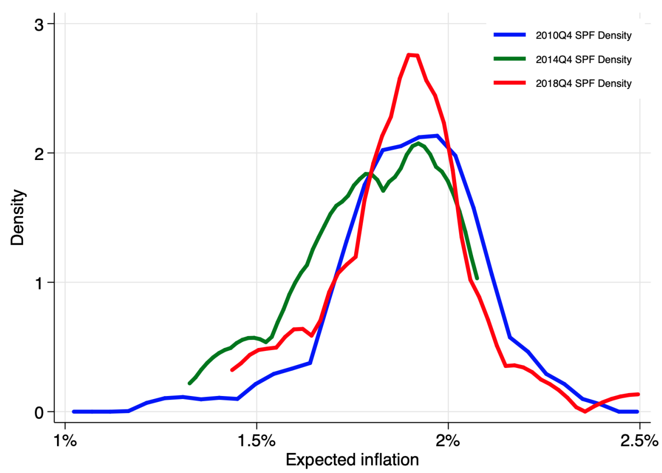
where π_t^e is the expected inflation that is the focus of this paper, and α is the vector with the sensitivity of inflation to each shock. Both are equilibrium objects, which the model will solve for.

Figure 9: The discrepancy and expectations in the Eurozone

(a) The people/traders in the SPF, and the markets



(b) The distribution of people/traders in the SPF



6.1 The transmission of the discrepancy to savings decisions

A central bank sets the nominal interest rate i_t^{CB} that prevails in financial markets. Yet, this rate is then intermediated through the financial markets, and absorbed through the beliefs of traders, and this transmission is captured by the discrepancy. The Euler equation, reflecting savings choices and the no-arbitrage relation between real and nominal rates, becomes:

$$g_t = \ln(\zeta) + i_t^{CB} + \alpha'\alpha - \pi_t^e - \delta\phi_t. \tag{15}$$

All the terms but the last are standard: the expected growth rate of consumption (and output) is equal to minus the subjective discount rate $(\ln(\zeta))$, plus the shadow real rate $(i_t^{CB} - \pi_t^e)$, plus the inflation risk premium $(\alpha'\alpha)$.[34]

The novelty is that a discrepancy in inflation expectations creates a wedge between the interest rate set by the central bank and that used by a representative agent. Intuitively, if the markets expected higher inflation than people (positive ϕ_t), then for a fixed policy nominal rate, the real interest rates offered by financial markets will be lower, and thus come associated with lower consumption growth. Monetary policy affects ϕ_t through its effects on expected inflation, and in turn ϕ_t is a proxy for the frictions in the transmission of monetary policy that create a wedge between the policy rates and the rates used for intertemporal decisions. There are surely other shocks and variables affecting this wedge, but this paper's focus is on the discrepancy.

A micro-foundation for this relation would rely on segmented markets. In a simple version, the policy rate is the one reflected in the price of bonds, so the shadow real interest rate out of financial markets is: $r_t = i_t^{CB} - \mathbb{E}^m(dp_t/dt)$. Households, within a period, live in an island isolated from the financial markets, from where they take as given the shadow real rate, and use their expectations of inflation to offer savings contracts at nominal rate: $i_t = r_t + \mathbb{E}^p(dp_t/dt)$. Finally, there is a representative agent making savings decisions, to which households and traders return at the end of the period, and who, by aggregating their information, has rational expectations over inflation. Combining all these ingredients returns equation (15) with $\delta = 1$.

More generally, I consider different values of $\delta \in [0,1]$ corresponding to different weights that the expectations of different agents have on ultimate savings decisions.

[34] A simplifying assumption left in the background is the link from the long-run inflation expectations discussed earlier to these expectations of short-run inflation. It would not be hard to extend the model to have bonds traded at different maturities and make this link. Since the model is affine, a standard affine term structure would follow leading to similar conclusions (Reis, 2019).

6.2 Monetary policy and its effects

The central bank sets interest rates according to a rule:

$$di_t^{CB} = -\rho(i_t^{CB} - i^*)dt + \eta\left(\frac{dp_t}{dt} - \pi^*\right) + \gamma d\phi_t. \tag{16}$$

The central bank has a long-run target for inflation, which coincides with the prior for inflation that agents use, and which is taken here to be constant. Keeping inflation and its expectations near to this target is the goal. Consistent with this target, nominal interest rates in the long-run are equal to $i^* = \pi^* + \ln(\zeta) + \alpha'\alpha$. The central bank smoothes interest rate adjustments towards that target, at a rate ρ, following a continuous-time partial adjustment process.

The two key policy parameters of interest to this paper are η and γ. The first describes how aggressive monetary policy is in response to inflation. It is well understood that η has to be sufficiently high to ensure that inflation expectations are anchored in the sense of delivering a determinate equilibrium for inflation. The question is whether that threshold—the Taylor principle—changes in the presence of the discrepancy.

The second parameter describes how the central bank responds to the discrepancy itself. Central banker speeches leave no doubt that policy responds to measures of inflation expectations, and the influence of the fundamental rational expectation of inflation is already present through the parameter η. The parameter γ captures the consequences of responding differentially to the different measures of expectations. In particular, the month-to-month changes in the discrepancy are mostly driven by asset prices, so that a positive γ corresponds to the central bank raising rates in response to market expectations of inflation rising.[35]

Monetary policy does not respond to output. This is a consequence of the final assumption: the classical dichotomy holds, so output growth will be exogenous to the model. This is consistent with the focus of the paper on long-run inflation expectations since, in the long run, arguably, prices are flexible and the Phillips curve is vertical.[36]

[35] Of course, ideally, the central bank would use the decomposition offered in this paper to separate disagreement within and across, obtain a measure of π_t^e and respond to that alone as opposed to the discrepancy. The question this section asks is instead positive, on how monetary policy responses to market expectations in the past have affected inflation anchoring.

[36] More practically, introducing firms and price stickiness would require introducing the expectations of firms distinct from those of households and markets, and other discrepancies, which are best left for future work. On the discrepancy between household and firm expectations, see Coibion, Gorodnichenko and Kumar (2018).

6.3 The discrepancy

Finally, the model of information of the previous sections translates into a function: $\phi_t = \Phi_t(\pi_t^e, \omega_t)$. Empirically, the last section effectively characterized this function for the US. In this section, to preserve the linearity of the model, I approximate this function log-linearly with:

$$\phi_t = \chi_\pi(\pi_t^e - \pi^*) + \chi_\omega \hat{\omega}_t, \tag{17}$$

where $\hat{\omega}_t = \log(\omega_t) - \log(1/2)$. I make a modest change to a previous assumption: $\hat{\omega}_t$ is normally distributed as opposed to having a (modified) Beta distribution, in order to keep the linear-Ito structure of the problem.

The discrepancy plays two roles in the economy. First, it affects monetary policy through the parameter γ, and it affects its transmission to inflation through the parameter δ. If both of these are zero, then we are back at the conventional analysis of inflation with an interest-rate rule. Second, the discrepancy is affected by the equilibrium expected inflation through the parameter χ_π, and it introduces financial shocks as drivers of inflation through the parameter χ_ω. Again, if both are zero, then the discrepancy would be zero, and we would get a conventional analysis.

6.4 Equilibrium

The model has two orthogonal shocks. The first are changes to output, which in this context are shocks to the natural rate of interest, or r-star. The second are the financial noise shocks. Both follow continuous-time autoregressive process:

$$dg_t = -\kappa_g(g_t - g^*)dt + \sigma_g dz_t^g \tag{18}$$

$$d\omega_t = -\kappa_\omega \hat{\omega}_t dt + \sigma_\omega dz_t^\omega. \tag{19}$$

All combined, an equilibrium of this monetary economy is a solution for inflation in (14) in terms of fundamental expected inflation π_t^e and the volatility of inflation in response to economic shocks α_g and noise shocks α_ω, subject to the Euler-Fisher equation in (15), the monetary policy rule in (16), and the discrepancy equilibrium law of motion (17), as well as the exogenous law of motion for the two shocks in equations (18)-(19).

6.5 Basics of anchoring: determinacy

A minimal definition of anchoring inflation is to make sure that the model has a determinate equilibrium. The appendix proves the following result:

Proposition 1. *Inflation is determinate as long as:*

$$\eta/\rho > 1 + \delta\chi_\pi \quad \textit{and} \quad \chi_\pi(\gamma - \delta) < 1. \tag{20}$$

If the discrepancy was exogenous, driven by noise and so unrelated to inflation, then $\chi_\pi = 0$ and the condition reduces to the standard Taylor principle: $\eta/\rho > 1$.[37] This would also be the case in the first condition if the discrepancy did not affect the transmission of policy rates to the savings decisions, so $\delta = 0$.

The presence of the discrepancy requires monetary policy to be unambiguously more aggressive in response to inflation. The reason is that when expected inflation rises, markets update by more than people, so the discrepancy rises. But, as a result, market interest rates fall, which pushes inflation up. This endogenous mechanism works against the response of policy to keep inflation determinate. Policy must therefore be more aggressive than before.[38]

6.6 Policy and the volatility of expected inflation

The appendix proves the following result:

Proposition 2. *Expected inflation equals:*

$$\pi_t^e = \pi^* + \frac{(\rho - \kappa_g)(g_t - g^*)}{\eta - \rho - \rho\delta\chi_\pi + \kappa_g(1 - \chi_\pi(\gamma - \delta))} + \frac{\chi_\omega[\kappa_\omega(\gamma - \delta) + \rho\delta]\hat{\omega}_t}{\eta - \rho - \rho\delta\chi_\pi + \kappa_\omega(1 - \chi_\pi(\gamma - \delta))}. \tag{21}$$

If inflation is determinate, then the conditions in proposition 1 ensure that the denominator in both fractions is strictly positive. The first term therefore reflects the standard result that an increase in the natural rate of interest (g_t), say because of a temporary fall in output, will tend to raise expected inflation, as long as interest rates persist relatively

[37] To see this is the standard Taylor principle, note that in the non-stochastic steady state, the interest rate rule in equation (16) is $i_t^{CB} = i^* + (\eta/\rho)(\pi_t^e - \pi^*)$.

[38] The second condition puts an upper bound on the policy response to the discrepancy γ. If that is far too strong, then the effect of inflation on the discrepancy leads to a policy change that raises inflation in the same direction, further pushes the discrepancy, and starts a process of indeterminacy.

little relative to the shock (so ρ is high relative to κ_g). More interesting, the higher is the response of policy to the discrepancy (γ), then the smaller is the volatility of expected inflation driven by natural-rate shocks. When the natural rate rises, this raises expected inflation, which raises the discrepancy as market prices reflect the shock faster than people. A central bank that is monitoring and responding to the discrepancy will then tighten more in response to these shocks, and thus lower the volatility of inflation.

As long as the discrepancy is contaminated by financial shocks ($\chi_\omega \neq 0$), then these shocks now also affect expected inflation. In one extreme, if monetary policy responds to the discrepancy ($\gamma > 0$), but it does not affect the transmission mechanism ($\delta = 0$), then a higher value of the discrepancy today (and so, given mean reversion, a future expected decline in it) lowers interest rates, which drives inflation up.[39] At the other extreme, if the discrepancy affects the transmission of monetary policy ($\delta > 0$), but leaves policy itself unchanged ($\gamma = 0$), then a positive financial shock raises market expectations of inflation, which lowers the real interest rate they offer in contracts to savers, which will raise inflation as long as interest rates persist relatively little relative to the shock (so $\rho > \kappa_\omega$). In between, when considering whether to respond to market prices and the discrepancy, policy must trade off the noise this brings to inflation, and the effects it has on the transmission of policy. In this particular simple model, the trade-off is such that choosing $\gamma = \delta - \delta\rho/\kappa_\omega$ would completely insulate expected inflation from the financial shocks.

However, there is a second trade off in keeping inflation expectations anchored. The discrepancy also provides a signal on the endogenous expected inflation itself ($\chi_\pi > 0$). Therefore, a stronger response to it by policy (higher γ) lowers the volatility of expected inflation coming from natural rate shocks. Monetary policy must weight the benefits of a stronger response to the discrepancy that takes advantage of its responsiveness to the natural-rate shocks to stabilize expected inflation against the costs that this response enhances the volatility of expected inflation sue to the financial shocks. Depending on the relative variance of the two shocks, an optimal γ that minimizes the unconditional variance of expected inflation, will be higher than $\delta - \delta\rho/\kappa_\omega$, depending on the relative variance of the two types of shocks.

To sum up, the optimal policy response to the discrepancy γ will be higher: (i) the higher is the direct effect of the discrepancy on the economy (δ), (ii) the less the discrep-

[39] Some algebra may clarify this statement: without shocks, the Taylor rule combined with the process for financial shocks becomes $di_t^{CB} = -\rho[i_t^{CB} - i^* - (\eta/\rho)\,(\pi_t^e - \pi^*) + (\gamma\chi_\omega\kappa_\omega/\rho)\hat{\omega}_t]dt + \gamma\chi_\pi d\phi_t$. So, a higher financial shock leads to looser policy.

ancy responds to financial shocks (χ_ω), (iii) the lower is the volatility of financial shocks (σ_ω), (iv) the more the discrepancy signals movements in expected inflation (χ_π), (iv) the higher the volatility of natural-rate shocks (σ_π). The benefits of responding to the discrepancy are two-fold: to offset the impact the discrepancy has on the transmission of interest rates to savings decisions, and to exploit an extra signal of changes in expected inflation in response to natural-rate shocks. The cost is that responding to the discrepancy transmits financial market shocks into movements in expected inflation.

6.7 Who is right: the people or the markets?

A well-known puzzle in the empirical expectations literature is that the surveys sometimes forecast one-year ahead US inflation better than markets (Ang, Bekaert and Wei, 2007). With the 5-year horizon, in the baseline sample, because of the bias in people's expectations, they have a higher mean-squared forecasts error than the market. However, demeaning both expectations series, the mean squared error of the market's forecasts is twice higher than that of the people, since markets have moved around a lot, but inflation has been quite steady.

The model can make sense of the mixed results from these horse races. In the model, because market traders are better informed, being able to observe prices, their expectations are more responsive to the fundamental expected inflation. At the same time, prices (and so traders' expectations) are contaminated by the noise that comes from the financial shocks. Which forecasts better will depend on whether households are minimally responsive to news (the size of the parameter θ in the model), and on whether markets are too responsive to financial noise (the size of χ_ω relative to χ_π).

Keeping these parameters fixed, monetary policy has a direct effect on this comparison. If monetary policy responds to the discrepancy (γ) in such a way that it keeps expected inflation insulated from financial shocks, then the trader's expectation may well forecast worse. Moreover, if monetary policy is so aggressive in responding to inflation (η) that it keeps expected inflation almost always on target, then even very uninformed people will be close to perfect forecasters (once removed from their experience bias), even as markets make errors due to financial noise. This may provide an approximation for the US results of the last decade.

There is a second, general-equilibrium, effect of monetary policy on the relative forecasting performance. The parsimonious model of people and market's expectations in this paper implies that policy will affect the updating of those expectations. To keep the discussion simpler, consider the case where the discrepancy does not affect interest rates directly ($\delta = 0$), and start from the case where monetary policy does not respond to the discrepancy ($\gamma = 0$). Imagine monetary policy announced it will become less responsive to inflation (η lower) to focus on some other objectives and this comes with a rise in inflation. Expected inflation will become more volatile. The people's expectations will likewise become more sensitive to the fundamental, but the overconfidence implies that this effect is small (θ increases slightly), while the stickiness of information and learning from experience imply that disagreement and skewness may rise as well. Altogether, people become worse forecasters. As for markets, since the fundamental has become more variable relative to the noise, asset prices become more informative on expected inflation. Endogenously, the size of χ_π relative to χ_ω rises. Relative to the people, markets become better forecasters.

Combining this information effect with the direct effect in the previous paragraph, an extended period of time where interest rates are held approximately unchanged irrespective of inflation, is a period where the discrepancy will become more volatile but also more informative. The benefits are larger from measuring the discrepancy, from decomposing it in its two disagreement terms, from using the model as a filter to estimate the underlying fundamental expected inflation, and from adjusting policy in response.

7 Conclusion

How are expectations of macroeconomic variables formed, and how should policy adapt to it? This paper added a new perspective on this classic question. It measured a new object to study the formation of expectations, the discrepancy between market prices and people's expectations of long-run inflation at a business cycle frequency. It proposed a parsimonious model of subjective expectations and financial markets that is flexible enough to fit the US survey data and market data. It offered an application to the US data, measuring the underlying fundamental expected inflation, and relating it to different dimensions of disagreement across groups of people, and within them. Finally, it laid out the trade-off that policy faces when choosing whether to respond to the discrepancy.

The paper reached a few conclusions. First, that the discrepancy in long-run inflation expectations has large business-cycle fluctuations, it is systematically related to monetary policy, and it is driven by disagreement across groups in the population as well as disagreement between the average and the marginal market traders. Second, that a combination of imperfect information, over-confidence, learning from experience, and sticky information, can explain the three first moments of the cross-sectional survey data on household long-run inflation expectations. Third, that the marginal trader in the US data was significantly bearish on inflation in 2014-2016, and this played large role in explaining the large negative discrepancy that arose during this time, as a consequence of this implied negative skew in traders' expectations. Fourth, after 2016, the clear reduction in the bias that generates a positive skew among households in the data lowered disagreement across households and traders, reducing the discrepancy while signaling a fall in fundamental expected inflation. Fifth, combining these two changes, long-run expected inflation has been trending down in the US since 2014 and is now around 1.8% (with the incomplete Eurozone data suggesting it may be significantly lower there). Sixth, according to the model, determinacy of inflation requires monetary policy to respond more aggressively to inflation. Seventh, monetary policy must trade off financial shocks versus natural rate shocks when choosing how much to respond to the discrepancy, and in doing so it may result in either markets or people being the better forecaster.

References

Abrahams, Michael, Tobias Adrian, Richard K. Crump, Emanuel Moench, and Rui Yu. 2016. "Decomposing Real and Nominal Yield Curves." *Journal of Monetary Economics*, 84(C): 182–200.

Afrouzi, Hassan, Spencer Yongwook Kwon, Augustin Landier, Yueran Ma, and David Thesmar. 2020. "Overreaction and Working Memory." *NBER Working Paper 27947*.

Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski. 2013. "A Theory of Asset Prices based on Heterogeneous Information." *CEPR Discussion Paper 9291*.

Andrade, Philippe, and Hervé Le Bihan. 2013. "Inattentive Professional Forecasters." *Journal of Monetary Economics*, 60(8): 967 – 982.

Andrade, Philippe, Richard K. Crump, Stefano Eusepi, and Emanuel Moench. 2016. "Fundamental Disagreement." *Journal of Monetary Economics*, 83(C): 106–128.

Andreasen, Martin M., Jens H. E. Christensen, and Simon Riddell. 2017. "The TIPS Liquidity Premium." *FRB San Francisco Working Paper 2017-11*.

Ang, Andrew, Geert Bekaert, and Min Wei. 2007. "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?" *Journal of Monetary Economics*, 54(4): 1163 – 1212.

Angeletos, George-Marios, Zhen Huo, and Karthik A. Sastry. 2020. "Imperfect Macroeconomic Expectations: Evidence and Theory." *NBER Macroeconomics Annual*, forthcoming.

Angeletos, G.-M., and C. Lian. 2016. "Incomplete Information in Macroeconomics." In *Handbook of Macroeconomics*. Vol. 2 of *Handbook of Macroeconomics*, , ed. J. B. Taylor and Harald Uhlig, Chapter 0, 1065–1240. Elsevier.

Bassetto, Marco, and Carlo Galli. 2019. "Is Inflation Default? The Role of Information in Debt Crises." *American Economic Review*, 109(10): 3556–84.

Beechey, Meredith J., Benjamin K. Johannsen, and Andrew T. Levin. 2011. "Are Long-Run Inflation Expectations Anchored More Firmly in the Euro Area Than in the United States?" *American Economic Journal: Macroeconomics*, 3(2): 104–29.

Binder, Carola, Wesley Janson, and Randal Verbrugge. 2019. "Thinking Outside the Box: Do SPF Respondents Have Anchored Inflation Expectations?" *FRB Cleveland Working Paper 19-15*.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*, 110(9): 2748–2782.

Cao, Shuo, Richard K. Crump, Stefano Eusepi, and Emanuel Moench. 2020. "Fundamental Disagreement about Monetary Policy and the Term Structure of Interest Rates." *CEPR Discussion Paper 15122*.

Carroll, Christopher D. 2003. "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics*, 118(1): 269–298.

Chernov, Mikhail, and Philippe Mueller. 2012. "The Term Structure of Inflation Expectations." *Journal of Financial Economics*, 106(2): 367–394.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics*, 115(1): 147–180.

Clements, Michael P. 2019. *Macroeconomic Survey Expectations.* Palgrave MacMillan.

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

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review*, 105(8): 2644–2678.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar. 2018. "The Formation of Expectations, Inflation, and the Phillips Curve." *Journal of Economic Literature*, 56(4): 1447–91.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. "How Do Firms Form Their Expectations? New Survey Evidence." *American Economic Review*, 108(9): 2671–2713.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal. 2009. "Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility." *Journal of Finance*, 64(2): 579–629.

Eusepi, Stefano, and Bruce Preston. 2018. "The Science of Monetary Policy: An Imperfect Knowledge Perspective." *Journal of Economic Literature*, 56(1): 3–59.

Eusepi, Stefano, Emanuel Moench, Bruce Preston, and Carlos Viana de Carvalho. 2019. "Anchored Inflation Expectations." *CEPR Discussion Paper 13900*.

Faust, Jon, and Jonathan H. Wright. 2013. "Forecasting Inflation." In *Handbook of Economic Forecasting*. Vol. 2 of *Handbook of Economic Forecasting*, , ed. Graham Elliott and Allan Timmermann, Chapter 1, 2 – 56. Elsevier.

Fleckenstein, Matthias, Francis A. Longstaff, and Hanno Lustig. 2014. "The TIPS-Treasury Bond Puzzle." *Journal of Finance*, 69(5): 2151–2197.

Galati, Gabriele, Richhild Moessner, and Maarten van Rooij. 2020. "The Anchoring of Long-term Inflation Expectations of Consumers: Insights from a New Survey." *De Nederlandsche Bank Working Paper 688*.

Garcia, Juan Angel, and Sebastian Werner. 2018. "Inflation News and Euro Area Inflation Expectations." International Monetary Fund IMF Working Papers 18/167.

Grossman, Sanford J, and Joseph E Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 70(3): 393–408.

Guo, Hongye, and Jessica A. Wachter. 2019. ""Superstitious" Investors." *NBER Working Paper 25603*.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2010. "The TIPS Yield Curve and Inflation Compensation." *American Economic Journal: Macroeconomics*, 2(1): 70–92.

Harris, Milton, and Artur Raviv. 1993. "Differences of Opinion Make a Horse Race." *The Review of Financial Studies*, 6(3): 473–506.

Harrison, J.Michael, and David M Kreps. 1979. "Martingales and arbitrage in multiperiod securities markets." *Journal of Economic Theory*, 20(3): 381 – 408.

Haubrich, Joseph, George Pennacchi, and Peter Ritchken. 2012. "Inflation Expectations, Real Rates, and Risk Premia: Evidence from Inflation Swaps." *The Review of Financial Studies*, 25(5): 1588–1629.

Hilscher, Jens, Alon Raviv, and Ricardo Reis. 2014. "Inflating Away the Public Debt? An Empirical Assessment." *NBER Working Paper 20339*.

Kose, Ayhan, Hideaki Matsuoka, Ugo Panizza, and Dana Vorisek. 2019. "Inflation Expectations: Review and Evidence." *CEPR discussion paper 13601*.

Malmendier, Ulrike, and Stefan Nagel. 2015. "Learning from Inflation Experiences." *Quarterly Journal of Economics*, 131(1): 53–87.

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

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. 2004. "Disagreement About Inflation Expectations." *NBER Macroeconomics Annual*, 18: 209–248.

Martin, Ian W. R., and Christian Wagner. 2019. "What Is the Expected Return on a Stock?" *Journal of Finance*, 74(4): 1887–1929.

Mertens, Elmar, and James M. Nason. 2018. "Inflation and Professional Forecast Dynamics: An Evaluation of Stickiness, Persistence, and Volatility." *BIS Working Paper 713*.

Milgrom, Paul, and Nancy Stokey. 1982. "Information, Trade and Common Knowledge." *Journal of Economic Theory*, 26(1): 17–27.

Nakamura, Emi, and Jón Steinsson. 2018. "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*." *Quarterly Journal of Economics*, 133(3): 1283–1330.

Reis, Ricardo. 2019. "Central Banks Going Long." In *Monetary Policy and Financial Stability: Transmission Mechanisms and Policy Implications*. , ed. Alvaro Aguirre, Markus Brunnermeier and Diego Saravia, 43–81. Central Bank of Chile.

Scheinkman, José A., and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy*, 111(6): 1183–1220.

Shefrin, Hersh. 2008. *A Behavioral Approach to Asset Pricing.* Academic Press.

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

Simsek, Alp. 2013. "Belief Disagreements and Collateral Constraints." *Econometrica*, 81(1): 1–53.

Veldkamp, Laura L. 2011. *Information Choice in Macroeconomics and Finance.* Princeton University Press.

Vives, Xavier. 2008. *Information and Learning in Markets: The impact of market microstructure.* Princeton University Press.

Woodford, Michael. 2003*a*. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S.* , ed. Joseph Stiglitz Phillipe Aghion, Romain Frydman and Michael Woodford, 25–58. Princeton University Press.

Woodford, Michael. 2003*b*. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton, NJ:Princeton University Press.

Xiong, Wei, and Hongjun Yan. 2010. "Heterogeneous Expectations and Bond Markets." *Review of Financial Studies*, 23(4): 1433–1466.

Yetman, James. 2020. "The Pass-through from Short-horizon to Long-horizon Inflation Expectations." *BIS Papers 111*.

Appendix – For Online Publication

Appendix A presents an alternative model of traders' beliefs, appendix B presents alternative estimates of fundamental expected inflation, while appendix C provides the proof of the propositions in section 6.

A Alternative model of traders' beliefs

Under the alternative model of trader's beliefs discussed at the start of section 4, let $H(.)$ be their prior distribution now. Then, their posterior is instead given by:

$$p(\pi_t^e|v^i, q) = \frac{g(q|\pi^e)f(v^i|\pi_t^e)h(\pi^e)}{\int g(q|\pi^e)f(v^i|\pi_t^e)h(\pi^e)d\pi^e}. \tag{A1}$$

In turn the solution for asset prices now is:

$$Q(v^*) = \frac{\int y(\pi^e)g(v^* - \pi^e)f(v^* - \pi^e)h(\pi_t^e)d\pi^e}{\int g(v^* - \pi^e)f(v^* - \pi^e)h(\pi_t^e)d\pi^e}. \tag{A2}$$

All the other parts of the model are unchanged.

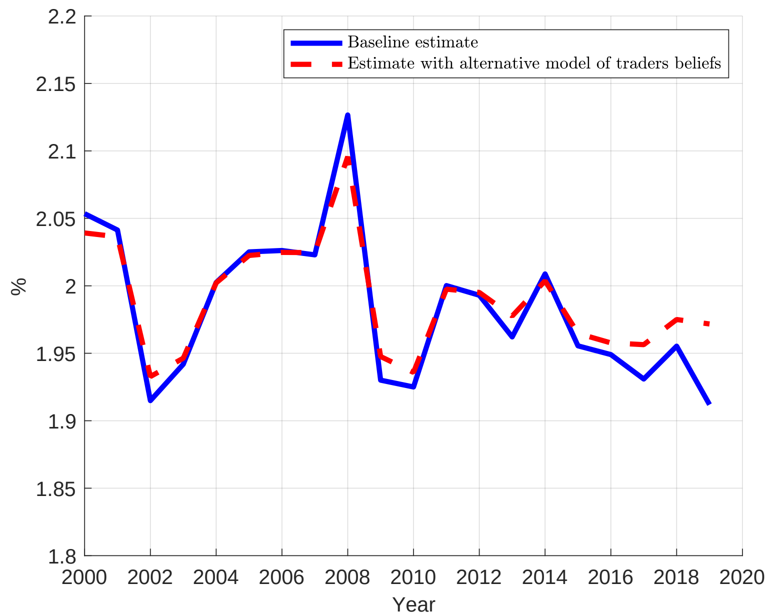
Figure A1 shows the median posterior estimates of π_t^e since 2000 using this alternative model. The main conclusion that expected long-run inflation moves around, so is not so well anchored, and has been steadily declining since 22014 remain.

B Alternative estimates of fundamental expected inflation

B.1 Dynamic prior

Figure A2 shows the estimates of the model where the parameter for the inflation target at each date is set to be equal to last year's posterior expected long-run inflation, as opposed to always equal to 2%. This makes the estimation procedure dynamic, and reflects the possibility that the private agents in the economy solve the same problem as the econometrician in this paper, and update what they believe the actual inflation target of the Federal Reserve is.

Figure A1: Expected long-run US inflation with alternative model of traders' beliefs

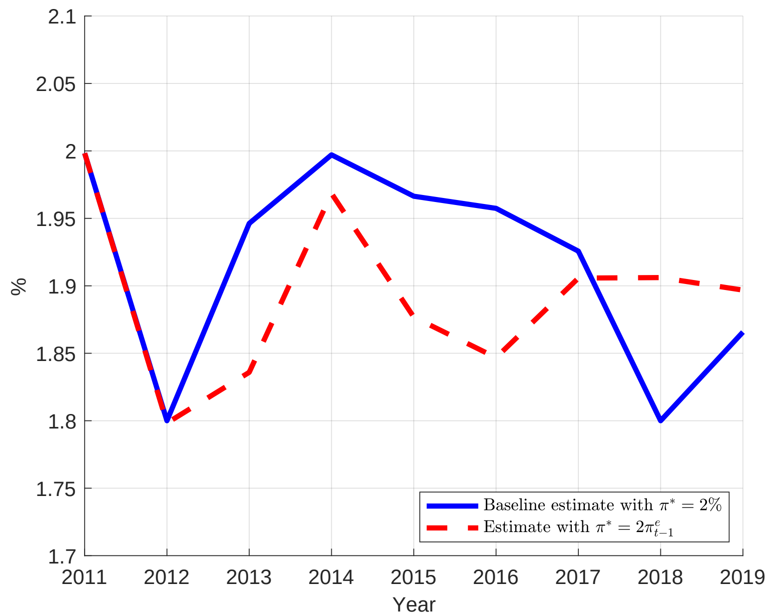


The pattern of a fall in expected long-run inflation since 2014 is still visible. But, as π^* now falls in 2014-16, the fall in the estimate of π^e in these years is more pronounced. At the same time, because most of the fall is realized by 2016, the data pointing to a new long-run anchor around 1.8 now leads to a slight rise in the estimates at the end of the sample.

B.2 Alternative choices of β

The parameter β determines how spread out the distribution of financial shocks is: if $\beta = 1$ then the distribution becomes uniform in the unit interval, whereas if $\beta \to \infty$, then the distribution becomes close to a point mass at 1/2. Figure A3 shows the robustness of the estimates to different choices of β. The top panel shows how the density changes for three choices of β, while the bottom panel re-estimates the model with these choices. Recall that the baseline in the text was $\beta = 2$. The inference that expected long-run inflation has been falling in the US seems robust.

Figure A2: Expected long-run US inflation with time-varying inflation target π^*



B.3 Monthly estimates

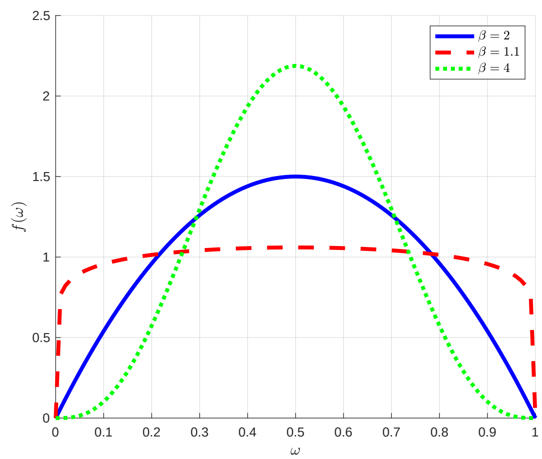
Figure A4 shows the estimates of the model solved at every month in the sample for which there is data on the dealers' expectation (8 times per year). For some months the numerical algorithm fails to find a π^e_t because the combination of asset prices and expectations would push for realized values of ω that are very close to the limits of 0 or 1. In those case, the figure sets the estimate to be the same as last month's. The estimates are very noisy, which is not surprising given the noise in the Michigan survey inputs observed in figure 5. Sitll the broad downward trend since 2014 can be seen.

B.4 The marginal trader since 2000

Figure A5 shows the median of the posterior distribution over who is the marginal trader in the market for inflation risk since 2000. The interesting new fact is the rise in the percentile of the marginal trader from 2002 to 2005. Much like the negative skewness among traders in 2014-16 pushed asset prices down, partly signaling a fall in long-run expected inflation, the same phenomenon in reverse was at play in 2002-05. However, unlike what

Figure A3: Robustness of estimates to choice of β

(a) Densities of ω for different choices of β



(b) Expected long-run US inflation for different choices of β

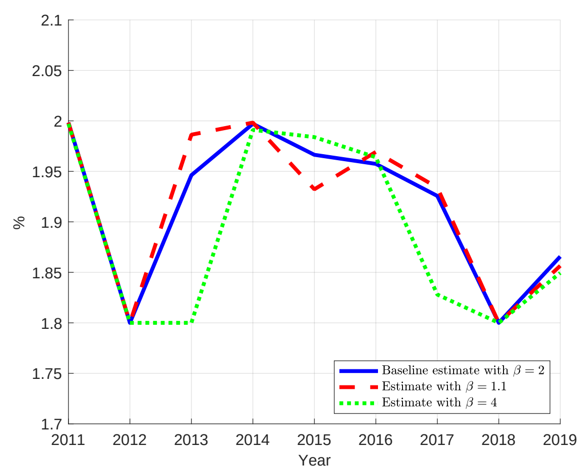
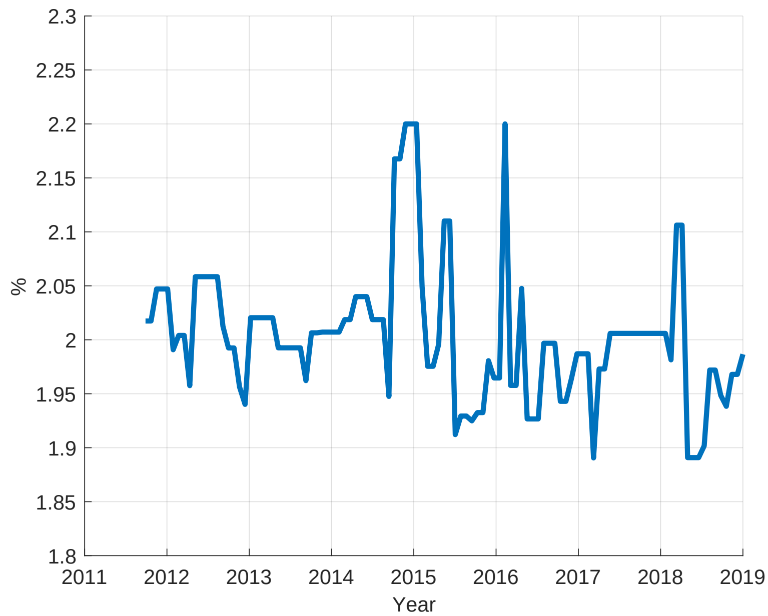


Figure A4: Expected long-run US inflation estimated every month



happened in 2017-19, the household survey never showed an, even delayed, move towards expecting higher inflation. Therefore, the model puts a small weight on this shift reflecting persistently higher long-run expected inflation, and reverses its estimates towards the mean soon after.

B.5 Alternative choices of the prior

The baseline post-2000 estimates used a normal prior for π_t^e centered at π^* with a standard deviation of 0.1, and a normal hyper-prior for π^* centered at 1% with standard deviation 0.1 Figure A6 shows the impact of raising the standard deviation of the first prior to 0.2. The median estimates are almost identical but, of course, the posterior credible sets are now wider.

Figure A5: The expected percentile of the marginal trader since 2000

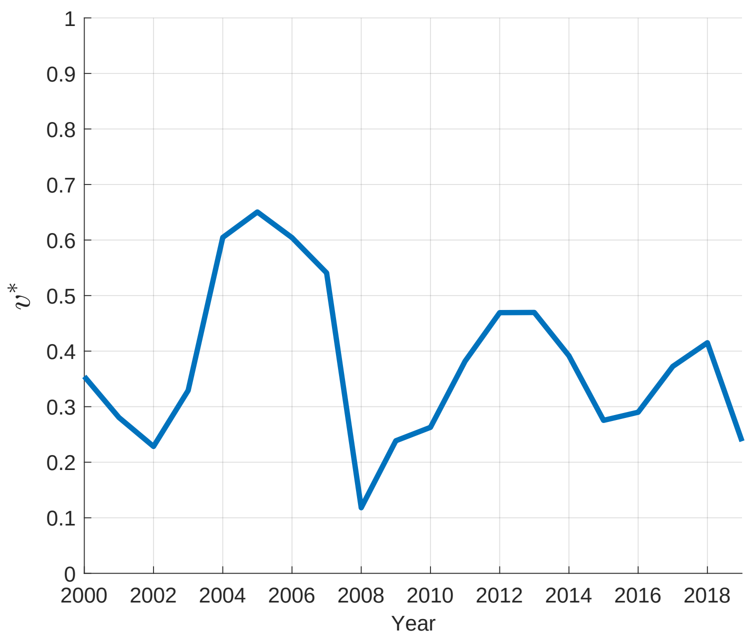
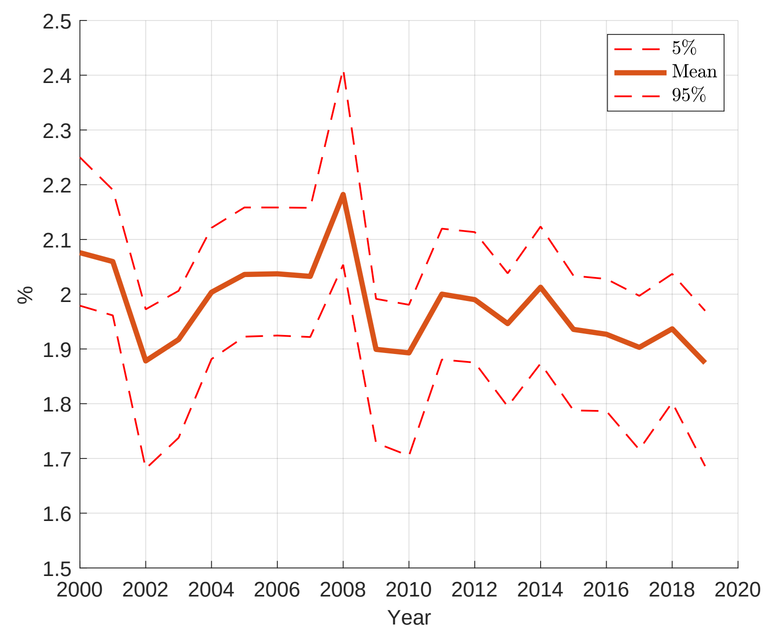


Figure A6: Expected long-run US inflation since 2000 with higher variance of prior



C Proof of propositions 1 and 2

Combining equations (14), (15), (16), and (17) to replace out dp/p_t, i_t^{CB}, and ϕ_t, gives the following long equation

$$dg_t + (1+\delta\chi_\pi)d\pi_t^e + \delta\chi_\omega d\omega_t = -\rho\left(g_t - g^* + \pi_t^e - \pi^* + \delta\chi_\pi(\pi_t^e - \pi^*) + \delta\chi_\omega\omega_t\right)dt \\ + \eta(\pi_t^e - \pi^*)dt + \eta\alpha\, dZ_t + \gamma\chi_\pi d\pi_t^e + \gamma\chi_\omega d\omega_t \quad \text{(A3)}$$

Following Reis (2019), take expectations at date t, recall the stochastic processes in equations (18)-(19), and rearrange to write this as:

$$(1+(\delta-\gamma)\chi_\pi)\,\mathbb{E}_t(d\pi_t^e) - (\eta - \rho - \rho\delta\chi_\pi)(\pi_t^e - \pi^*)dt = \\ (\kappa_g(g_t - g^*) + (\delta-\gamma)\chi_\omega\kappa_\omega\omega_t)dt - \rho\left(g_t - g^* + \delta\chi_\omega\omega_t\right)dt. \quad \text{(A4)}$$

Imposing the usual terminal condition that expected inflation must stay bounded, then this equation will have a unique bounded solution only if:

$$\frac{\eta - \rho - \rho\delta\chi_\pi}{1+(\delta-\gamma)\chi_\pi} > 0. \quad \text{(A5)}$$

This is the result in proposition 1.

Solving the ordinary differential equation delivers:

$$\pi_t^e = \pi^* + \int_t^\infty e^{-\left(\frac{\eta-\rho-\rho\delta\chi_\pi}{1+(\delta-\gamma)\chi_\pi}\right)s}\left[\frac{(\rho-\kappa_g)\,\mathbb{E}_t(g_s - g^*) + \chi_\omega(\rho\delta + (\gamma-\delta)\kappa_\omega)\,\mathbb{E}_t(\omega_s)}{1+(\delta-\gamma)\chi_\pi}\right]ds. \quad \text{(A6)}$$

Using the properties of Orstein-Uhlenbeck processes to take the expectations, evaluating the integral, and rearranging gives the result in proposition 2.