Federal Reserve Bank of Cleveland Working Paper Series

Aggregate Implications of Heterogeneous Inflation Expectations: The Role of Individual Experience

Mathieu O. Pedemonte, Hiroshi Toma, and Esteban Verdugo

Working Paper No. 23-04

January 2023

Suggested citation: Pedemonte, Mathieu O., Hiroshi Toma, and Esteban Verdugo. 2023. "Aggregate Implications of Heterogeneous Inflation Expectations: The Role of Individual Experience." Working Paper No. 23-04. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202304.

.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Aggregate Implications of Heterogeneous Inflation Expectations: The Role of Individual Experience*

Mathieu Pedemonte
Federal Reserve Bank of Cleveland

Hiroshi Toma
University of Michigan

Esteban Verdugo
University of Michigan

January 2, 2023

Abstract

We show that inflation expectations are heterogeneous and depend on past individual experiences. We propose a diagnostic expectations-augmented Kalman filter to represent consumers' heterogeneous inflation expectations-formation process, where heterogeneity comes from an anchoring-to-the-past mechanism. We estimate the diagnosticity parameter that governs the inflation expectations-formation process and show that the model can replicate systematic differences in inflation expectations across cohorts in the US. We introduce this mechanism into a New Keynesian model and find that heterogeneous expectations anchor aggregate responses to the agents' memory, making shocks more persistent. Central banks should be more active to prevent agents from remembering current shocks far into the future.

Keywords: Expectations, survey data, belief formation, heterogeneous expectations
JEL codes: D84, E31, E58, E71

*We thank Andres Blanco, Olivier Coibion, Javier Cravino, Ina Hajdini, Christopher House, Daniel John, Ed Knotek, John Leahy, Jean-Paul L'Huillier, Matthew Shapiro, Michael Weber, Donghoon Yoo, Basit Zafar and the seminar participants at the BSE Workshop on Expectations in Macroeconomics, the KIEP-CEPR Return of Inflation Conference, the Federal Reserve Bank of Cleveland and the University of Michigan for their useful comments. The views expressed herein are those of the individual authors and do not necessarily reflect the official positions of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

1 Introduction

Inflation expectations matter for decisions at both the firm and the household levels (Coibion et al. 2019; Coibion, Gorodnichenko, and Ropele 2020; Hajdini et al. 2022b). Given their importance, there is an increased interest in measuring them and exploring what determines their formation process. The recent literature shows that individuals form inflation expectations, for instance, based on their recent buying experience (D'Acunto et al. 2021) and historical experiences regarding aggregate inflation (Malmendier and Nagel 2016). While most studies on this topic provide empirical evidence regarding how differences in inflation expectations arise at the individual level and their effects on different micro-level decisions, there is less understanding of the aggregate implications associated with the heterogeneity of inflation expectations observed in the data. There is a noticeable gap in the literature between the empirical micro-level findings and macroeconomic models.

This paper aims to fill this gap. First, we show that individuals' inflation expectations depend on their history of inflation, confirming the main empirical finding of Malmendier and Nagel (2016) but using new US and international evidence.[1] Using detailed micro-level data, we find that (i) inflation expectations are heterogeneous across cohorts, (ii) inflation experiences are clustered by age and (iii) positively correlated with individual inflation expectations, and (iv) there are no differences between cohorts in updating to current information.

Based on these facts, we propose a diagnostic expectations-augmented Kalman filter (Bordalo, Gennaioli, and Shleifer 2018; Bordalo et al. 2019, 2020) to model the inflation expectation-formation process. We depart from rational expectations by presenting a formulation where individuals under-react to recent developments in the inflation rate, by positively weighting their inflation histories.[2] As a result, agents who experienced episodes of high inflation in the past systematically forecast higher inflation values compared to those who experienced episodes of low inflation.[3] Thus, our framework predicts patterns that match the data.

Under the proposed framework, inflation expectations have two components: a current forecast made with shared current information between agents and a reference term that depends on individuals' past experiences. We structurally estimate the diagnostic parameter that governs the expectations-formation process. As all respondents share current information, we can control for the current common forecast using time fixed effects. The resulting estimation for the diagnostic coefficient is -0.324, implying that consumers under-react to recent

[1] While Malmendier and Nagel (2016) use the University of Michigan's Survey of Consumers (MSC), in Sections 3 and 4 we provide external validity to their result by using a different data source with panel characteristics: the Survey of Consumer Expectations (SCE) from the Federal Reserve Bank of New York. In addition, in Appendix D we show that the finding is consistent at the international level even after controlling for common cohort characteristics.

[2] This under-reaction effect is a conclusion obtained after estimating the key diagnostic parameter using detailed micro-level data.

[3] Notice that this result does not imply that agents cannot over-react to some news. Our empirical exercise shows that the proposed expectation modeling framework explains the heterogeneity observed across cohorts, but the presence of a common component in the modeling allows for a common over-reaction to some current news.

news on inflation, positively weighting their inflation history. This experience-based bias can explain the inflation expectations heterogeneity across cohorts observed in surveys.

These two modeling components of expectations effectively account for a relevant part of the heterogeneity observed in the data. First, we model the shared component following a standard signal-extraction procedure. Based on works exploring reference prices for consumers (for example, D'Acunto et al. 2021), we use the lagged inflation rate of the food component of the CPI as a signal of the non-observed aggregate inflation variable. Second, we construct a cohort-specific inflation expectations measure using (i) the already computed shared component, (ii) an idiosyncratic element related to the memory of the cohort, and (iii) the diagnostic coefficient value estimated in the empirical section. As a result, we obtain model-based forecasts that closely match the heterogeneous inflation expectations across cohorts and time observed in the data. A regression between our diagnostic model-based cohort-specific inflation expectations and the expectations we see in the survey data gives a coefficient of 0.899, statistically different from zero at the 99 percent confidence level. Although the model-based inflation forecasts do not use micro-level information about inflation expectations, they remarkably predict survey data. Moreover, while our focus is on cohort heterogeneity, the common component across agents also shows good properties for modeling inflation expectations using signals from grocery prices.

We then introduce this heterogeneous inflation expectations mechanism into a New Keynesian model to explore the macro implications arising from this micro-level heterogeneity. We allow households to form expectations according to the proposed diagnostic Kalman filter. In our model, while old generations have their expectations shaped mainly by their past, new generations are highly influenced by recent developments. We find that heterogeneous expectations anchor the aggregate response of the inflation and output gap to agents' memory. At the same time, they also increase the duration of the effects of the shocks.[4]

With this setting, extrapolation works differently than the over-reaction found in L'Huillier, Singh, and Yoo (2021). Starting from the steady state and after a negative supply shock, consumers are over-optimistic, as their expectations are anchored to the steady state, especially for older cohorts. This effect reduces the decline in output. Rational firms translate the cost shock to higher prices, making the effect less negative in terms of output. The opposite happens after a positive demand shock, where consumers are over-pessimistic regarding the output gap but optimistic about inflation. These gaps reduce the pace of price adjustment, creating a smaller but very persistent inflation.

We then perform an optimal Taylor rule exercise where the central bank seeks to minimize the expected volatility of the economy by optimally choosing the parameters of the Taylor rule. When we allow for heterogeneous expectations in the model, agents have long memories and remember current shocks far into the future. After a negative supply shock or a positive demand shock, the optimal response of the central bank is to be more active. This way, the

[4]While we could incorporate other forms of heterogeneity and biases related to experience-based mechanisms on the firm or the government side of the economy, in this paper, we introduce non-rational heterogeneous expectations only on the household side. We make this decision because data that combines information about expectations and age is available only for households.

monetary authority prevents inflation from rising and, more importantly, prevents agents from incorporating a high-inflation episode into their memories.

This paper has important implications for explaining past inflation dynamics and learning about the future consequences of recent developments for the economy. Since 2021, a new cohort of consumers worldwide has been experiencing relatively high inflation rates for the first time in their lives. According to our findings, this high-inflation episode could have consequences in the middle run since consumers incorporate this episode into their history of inflation, adjusting future expectations. Our framework shows that permitting high inflation produces higher and more persistent inflation expectations, which in turn generate a higher and more persistent inflation rate in the future. Our findings help us understand why inflation has persisted in the past, why consumers' inflation expectations are persistent today, what to expect from episodes of unusually high inflation, and how central banks should react to such episodes.

Recent macroeconomic models show the relevance of heterogeneity in explaining aggregate fluctuations. However, the focus has been mostly on household financial constraints (Kaplan, Moll, and Violante 2018), and there is little evidence on the role of heterogeneity of expectations across individuals. Although surveys show significant heterogeneity across firms (Coibion, Gorodnichenko, and Kumar 2018) and household inflation expectations (Hajdini et al. 2022c), few works study its macroeconomic implications. A notable exception is Afrouzi (2020), who shows that heterogeneity in firm-level inflation expectations, coming from different levels of attention due to information acquisition about competitors' beliefs, amplifies monetary non-neutrality. Our paper focuses on the heterogeneity of expectations on the household side of the economy. Heterogeneous inflation expectations anchor the response of aggregate variables to agents' memory, increasing the persistence of the effects of shocks. Therefore, an energetic reaction from monetary authorities prevents current high inflation and prevents agents from incorporating high-inflation episodes into their memories, thus preventing higher future inflation expectations.

In addition, other works focus on exploring the source of heterogeneity across firms and consumers (for example, Afrouzi 2020; Hajdini et al. 2022c). Malmendier and Nagel (2016) find that personal experience affects individuals' expectations formation. D'Acunto et al. (2021) show that personal buying experiences matter. These experiences create systematic differences in expectations across the population. In addition, Coibion et al. (2019), Roth and Wohlfart (2020) and Hajdini et al. (2022b) show that consumer expectations matter for their decision-making process.

The rest of the paper is organized as follows. Section 2 discusses recent works on the topic. Section 3 provides empirical results regarding heterogeneity in inflation expectations. We empirically model inflation expectations depending on the history of inflation experienced by cohorts in Section 4. Section 5 discusses the aggregate implications arising from heterogeneous inflation expectations. Section 6 shows results for an optimal Taylor rule exercise. We then analyze the high-inflation episode of 2021 through the lens of our theoretical model in Section 7. Finally, Section 8 concludes.

2 Literature review

A vast empirical literature shows that households' inflation expectations depart from full information rational expectations and are heterogeneous. Relevant to our paper, Malmendier and Nagel (2016) document that households present learning from past inflation mechanisms when forming inflation expectations. Thus, people who have experienced higher inflation rates in the past have higher inflation expectations for the future.[5] Therefore, heterogeneity of inflation expectations naturally arises as a result of different experiences with past inflation rates. Similar results are discussed in Malmendier (2021) and Malmendier and Wachter (2022).[6]

The relevance of inflation expectations at the household level is that they affect a broad set of household decisions. For instance, Malmendier and Nagel (2016) show that inflation expectations influence individuals' financial decisions, while Coibion et al. (2019) conclude that inflation expectations affect households' spending on durable goods.

However, households' inflation expectations are not only dependent on past experiences. Evidence shows they also respond to other variables such as professional forecasts (Carroll 2003), updating costs (Branch 2004), prices exposure (D'Acunto et al. 2021), and socioeconomic characteristics (D'Acunto, Malmendier, and Weber 2022; Madeira and Zafar 2015; Souleles 2004), among others. To further depart from the rational expectations hypothesis, there is evidence showing that household inflation expectations have a zero lower bound (Gorodnichenko and Sergeyev 2021) and disagree with financial markets (Reis 2020).[7]

The theoretical model proposed in this paper features monetary policy, a learning from the past mechanism, and overlapping generations. This framework connects this paper to several strands of the theoretical literature. In the first place, by departing from the rational expectations paradigm, our theoretical model inserts itself in to the literature of behavioral New Keynesian models such as the ones proposed by Gabaix (2020) and Jump and Levine (2019). Other references to New Keynesian and monetary models where agents learn instead of following rational expectations include Adam (2003), Airaudo, Nisticò, and Zanna (2015), Brazier et al. (2008), Evans, Honkapohja, and Marimon (2001), Gáti (2020), Jump, Hommes, and Levine (2019) and Orphanides and Williams (2007). Our model is also related to the New

[5]Similarly, Malmendier, Nagel, and Yan (2021) show that this learning from the past mechanism is also present in members of the Federal Open Market Committee (FOMC).

[6]In the literature, this learning from the past mechanism is not only associated with inflation expectations. Kuchler and Zafar (2019) show that individual expectations of housing prices and aggregate unemployment are also affected by past experiences. Additionally, Ehling, Graniero, and Heyerdahl-Larsen (2018); Nagel and Xu (2022); Malmendier, Pouzo, and Vanasco (2020a) and Malmendier, Pouzo, and Vanasco (2020b) show that this mechanism is also relevant in financial markets. While not directly referenced as learning from past experiences, a similar mechanism in which agents in financial markets learn from past realizations of the data appears in Collin-Dufresne, Johannes, and Lochstoer (2017) and Kozlowski, Veldkamp, and Venkateswaran (2020).

[7]While we focus on the household side of the economy, there is also evidence showing that professional forecasters depart from rational expectations too (for example, Bordalo et al. (2020); Coibion and Gorodnichenko (2015); Gáti (2020)). Because of data availability, evidence on the firm side is notably scarce (see Candia, Coibion, and Gorodnichenko (2022)).

Keynesian models in which there are different agents with heterogeneous expectations among them. Papers that analyze this situation include Branch and McGough (2009), Branch and McGough (2010), Branch and Evans (2011), De Grauwe (2011), De Grauwe and Ji (2020), Pfajfar (2013), and Massaro (2013). Additionally, we connect to papers such as Branch and McGough (2009), Di Bartolomeo, Di Pietro, and Giannini (2016) and Gasteiger (2014) by studying optimal monetary policy in a heterogeneous expectations context. These papers, however, do not consider heterogeneity stemming from having different cohorts that have gone through different experiences in the past.

By introducing overlapping generations in a New Keynesian context, we relate to Carton (2012), Gali (2021), and Fujiwara and Teranishi (2008). Furthermore, by stating that cohorts show heterogeneity in expectations because of different experiences lived in the past, we connect to papers unrelated to monetary policy but where different cohorts have different beliefs about the future, such as Collin-Dufresne, Johannes, and Lochstoer (2017), Malmendier, Pouzo, and Vanasco (2020b), and Schraeder (2016).

The framework we use to model non-rational and heterogeneous expectations is based on diagnostic expectations, as introduced for instance in Bordalo, Gennaioli, and Shleifer (2018), Bordalo et al. (2019), and Bordalo et al. (2020).[8] This framework has recently been applied to macro monetary settings in Bianchi, Ilut, and Saijo (2021), and L'Huillier, Singh, and Yoo (2021).

Besides diagnostic expectations, there is a large literature that focuses on ways of departing from the full information rational expectations assumption and analyzing the aggregate implications. Examples include the imperfect information approach (Mankiw and Reis 2011), the complex systems/animal spirits/heuristic approach (Branch and McGough 2010, 2018; Brock and Hommes 1997; De Grauwe 2011; De Grauwe and Ji 2020; Hommes 2021), the sticky information approach (Coibion 2006; Mankiw and Reis 2002, 2007; Reis 2009), the adaptive learning approach (Evans and Honkapohja 2001; Marcet and Sargent 1989), among others. However, few of these papers have studied how heterogeneity in expectations arises and its macroeconomic implications.

Our findings are closely related to those of Malmendier and Nagel (2016), who use an adaptive learning approach to approximate cohorts' heterogeneous inflation expectations. Instead, we opt for a diagnostic Kalman filter. While we also rely on constant gain, in our framework the selection of parameters is primarily data-driven. Our approach also allows us to incorporate a current shared forecast component using a standard Kalman filter and a structure that incorporates past inflation experiences.

Moreover, as we show later in the paper, we find that cohorts do not adjust their expectations differently in response to current inflation news. Our results suggest that while younger cohorts implicitly put more weight on current information, they adjust to the news data in the same way as older cohorts do. Younger cohorts do not react more strongly to inflation

[8] Our approach is also related to broader signal extraction/noisy information approaches such as the ones proposed in Blanchard, L'Huillier, and Lorenzoni (2013), Hürtgen (2014), Nimark (2014), and Woodford (2001).

news than older cohorts. In that sense, our modeling takes the approach of Malmendier and Nagel (2016), incorporating a method consistent with under-reaction to past experiences but allowing agents to use the new information to form expectations. In addition, our approach is flexible enough to be incorporated into a general equilibrium framework following the recent work of Bianchi, Ilut, and Saijo (2021) and L'Huillier, Singh, and Yoo (2021).

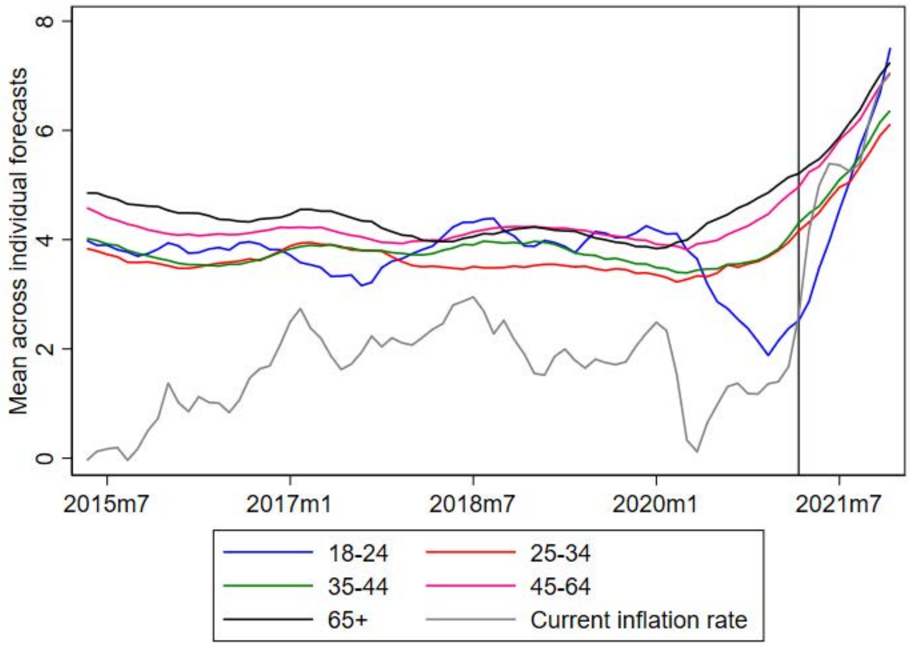
3 Empirical facts

In this section, we review some empirical facts related to heterogeneous inflation expectations on the household side and show how they are correlated with past experiences with inflation. These empirical facts help motivate and guide the theoretical model of the paper.

It has been documented that consumers' experiences influence their inflation expectations (D'Acunto, Malmendier, and Weber 2022; Malmendier 2021). This means that individual experiences are a source of heterogeneity in expectations. In this paper we focus on how inflation experiences influence inflation expectations, as in Malmendier and Nagel (2016). In the US, this heterogeneity has turned out to be particularly important after the high-inflation episode of 2021, when inflation surged after 30 years of low and stable rates, meaning that a new cohort of consumers might potentially be influenced by this event.

For this section we use data from the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. This is a US-wide rotating panel with information from March 2013 to December 2021. Each respondent is surveyed for a maximum of 12 contiguous months. This data-set is especially useful for our purposes because it provides high-frequency data on inflation expectations for American households in two different periods of the US economy. In particular, we focus on respondents' 12-months-ahead point forecast. The 12-months-ahead inflation rate is calculated as the inflation rate between the current month and 12 months after the current month.

Figure 1: Average 12-months-ahead inflation expectations



Note: Figure shows the 12-month moving average for the 10 percent and 90 percent trimmed mean for each cohort using the point forecast. We use population weights. Data go from June 2013 to December 2021. Ages correspond to the interviewee's age at the time of the survey. Vertical line denotes March 2021.
Source: Survey of Consumer Expectations.

Fact 1: Inflation expectations are heterogeneous across cohorts

Figure 1 shows the mean 12-months-ahead inflation forecast by cohort. The heterogeneity across cohorts is clear. The oldest cohort (65+) and the second oldest (45-64) have higher mean inflation expectations throughout most of the sample. Those cohorts experienced the period of high inflation in the 60s, 70s, and early 80s. They are followed by the intermediate cohorts (25-34 and 35-44), who experienced the stable and low inflation rates of the 90s, 00s, and 10s.

The youngest cohort (18-24) shows the most volatile mean, following the current inflation rate most of the time. For instance, their mean increases after the high-inflation episode of 2021, even surpassing the expectations of older cohorts.

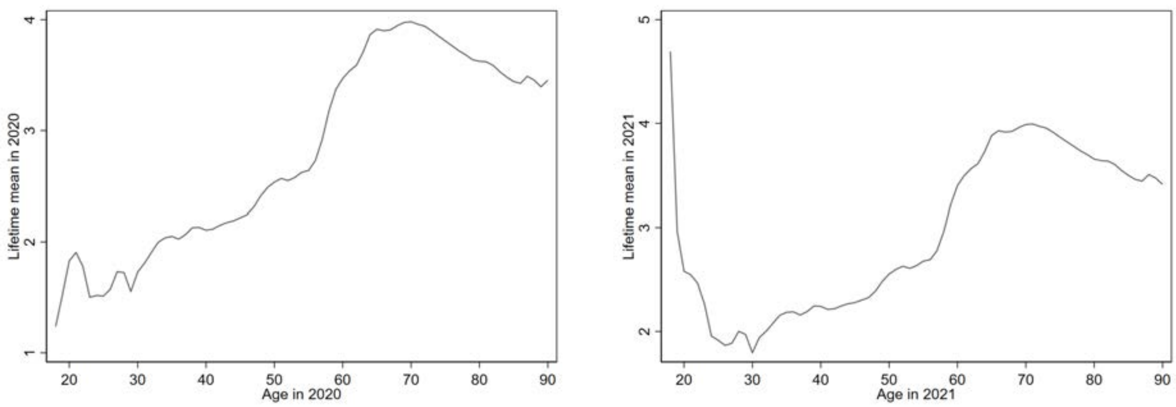
Fact 2: Inflation experiences are clustered by age

Figure 2 shows the average lifetime inflation rate people have experienced according to their age in 2020 and 2021. We see that, in the US, average lifetime inflation rates are clustered by age.

The heterogeneity of average experienced inflation rates across cohorts is a result of the different inflation-related events Americans have gone through. The older cohorts have experienced events such as the Great Inflation period (1965-1982), which was characterized by a high and persistent inflation rate, and thus have a higher lifetime average inflation rate, no matter for which year we make the calculations (2020 or 2021). Meanwhile, the intermediate cohorts have experienced periods of low and stable inflation rates throughout the 80s, 90s, 00s and 10s, for which they present a low lifetime average inflation rate. For older and intermediate cohorts, experiencing the high-inflation episode of 2021 did not affect their lifetime average inflation rate by much.

In contrast, we can see that the yougest cohorts do present a change in between the years 2020 and 2021. Up to 2020, the youngest cohorts had not experienced high inflation and, thus, show a low lifetime average inflation rate. However, after being exposed to the high-inflation episode of 2021, their lifetime average inflation rate dramatically increases.

Figure 2: Lifetime average inflation rate among respondents



(a) 2020 (b) 2021

Note: Figure shows the mean of the monthly YoY inflation rate that people of the age shown in the years 2020 and 2021 have experienced in their lifetimes, starting when they were age 18.
Source: FRED.

Fact 3: A higher average lifetime inflation rate is correlated with a higher point forecast

Tying together the previous two empirical facts, Figure 3 shows that people who have experienced higher average inflation rates during their lifetimes, when surveyed, tend to give higher point forecast for inflation.[9]

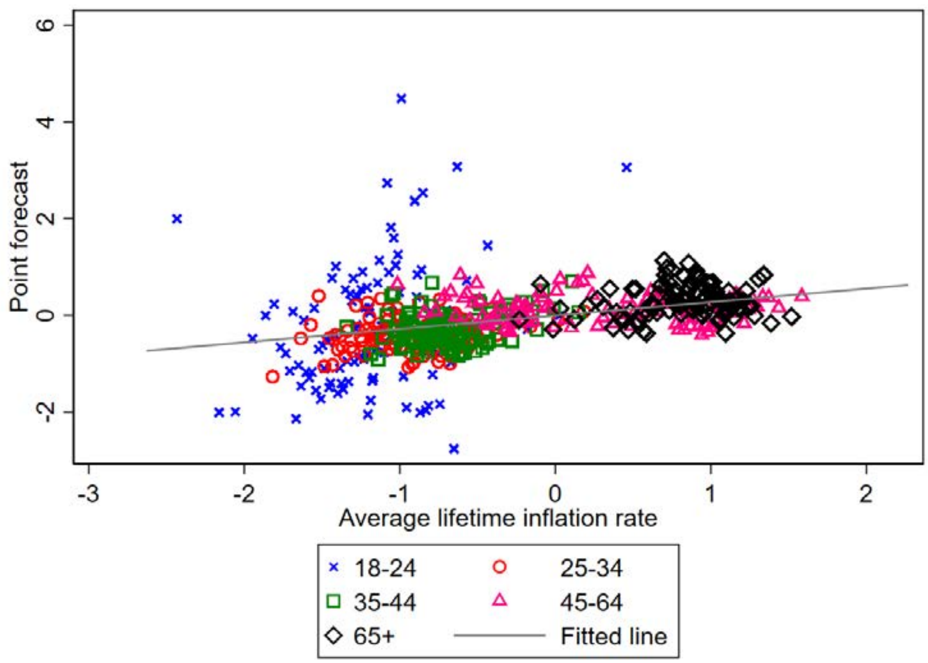
We formally test this result in Table 1. Columns 3 and 4 show that the inflation experienced has a significant effect on individuals' inflation expectations, after controlling for the current environment and individual characteristics.

This fact provides empirical support for the literature on learning from past experiences (Malmendier and Nagel 2016; Malmendier 2021; Malmendier, Nagel, and Yan 2021; Malmendier and Wachter 2022) and points to a possible source of heterogeneity in inflation expectations: past experiences with inflation.

[9]We control for observable characteristics of the respondent with the exception of age and period.

While the evidence we provide here is for the US, Hajdini et al. (2022a) find similar evidence in a survey for a panel of countries. They find that average inflation experience is positively related to individual inflation expectations, even after controlling for country-time fixed effects and, more importantly, cohort fixed effects that control for the fact that cohorts can have biases because of their age. More importantly, in the panel of countries they use, there are different inflation experiences across countries, which are not necessarily similar to those of the US. In addition, in Appendix D we find similar evidence for a panel of European countries, showing that this pattern (i) is also present beyond the US and (ii) does not arise from systematic characteristics of the cohorts, but because of the cross-country heterogeneous inflation experiences.

Figure 3: Inflation point forecast and average lifetime inflation



Note: Figure shows binned scatterplot across lifetime average inflation bins. Variables residualized by respondent gender and commuting zone. Data go from June 2013 to December 2021. Ages correspond to the interviewee's age at the time of the survey.
Source: Survey of Consumer Expectations.

Table 1: Effects of current and experienced inflation on inflation expectations

Dep. var.: Inflation expectations	(1)	(2)	(3)	(4)
Average lifetime inflation	0.332***	0.269***	0.299***	0.250***
	(0.026)	(0.065)	(0.025)	(0.022)
Current inflation	0.524***	0.632***		
	(0.016)	(0.118)		
Cohort 25-34		-0.080		
		(0.310)		
Cohort 35-44		-0.131		
		(0.302)		
Cohort 45-64		-0.010		
		(0.319)		
Cohort 65+		0.062		
		(0.325)		
Current inflation × 25-34		-0.163		
		(0.128)		
Current inflation × 35-44		-0.099		
		(0.126)		
Current inflation × 45-64		-0.080		
		(0.122)		
Current inflation × 65+		-0.141		
		(0.126)		
Time FE	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	105,415	105,415	105,415	105,402
R-squared	0.057	0.058	0.091	0.198

Note: Table shows regressions where the dependent variable is inflation expectations according to the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. Column (1) shows controls by the average lifetime inflation of respondents of a given age at each period of time and the last inflation measure. Column (2) follows (1) but adds cohort fixed effects and the interaction of those cohort fixed effects with the current inflation. Column (3) follows (1) but adds time fixed effects and, hence, omits the current inflation variable. Column (4) follows (1) but adds time fixed effects and demographic controls. The demographic controls are income, gender, Hispanic origin, race, educational level, numerical proficiency, and commuting zone. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by age. The dependent variable is trimmed, dropping the lower and upper 10 percent of answers in each period.

Fact 4: After controlling for the average lifetime inflation rate, younger cohorts do not react more strongly to inflation news than older cohorts

We also test whether younger generations react more strongly to the current economic environment, after controlling for their average lifetime inflation. This way we check for the results of Malmendier and Nagel (2016), where younger cohorts, because they have fewer observations and are learning about the economy, react more strongly to current events when compared to older cohorts.

We test this through a set of regressions at the individual level in Table 1. Column 1 shows that all individuals do react to current inflation events, as has been shown in other papers with information treatment (for example, Hajdini et al. 2022b). These results also confirm the existence of a positive relationship between the inflation forecasts and average lifetime inflation rates, as we saw in Figure 3, even after considering current inflation.

Then, to test whether there are differentiated reactions across cohorts, we run regressions that consider an interaction of current inflation with an indicator variable by cohort. Column 2 in Table 1 shows the results. We see that, after controlling for average lifetime inflation, the interaction term does not have a statistically significant effect, such that there are no different reactions to current inflation news across different cohorts, which contrasts with the finding of younger cohorts reacting more strongly to current events from Malmendier and Nagel (2016). This result suggests that the heterogeneity across cohorts comes from the different past experiences with inflation.

4 A simple model with heterogeneous expectations

In this section, we propose a diagnostic expectations-augmented Kalman filter (Bordalo et al. 2019) as the process by which agents form their inflation expectations. We begin with a simple model that provides a good starting point from which differences in agents' personal experiences do not imply heterogeneity in expectations. Given the absence of private information, we show that the observed heterogeneity cannot arise from a standard Kalman filter. Then, by introducing a diagnostic expectations-augmented Kalman filter, we explain how the inflation history experienced by the agents distorts their expectations, generating some heterogeneity in their inflation forecasts. Moreover, we estimate the corresponding distortion parameter. Lastly, we close the section by comparing these heterogeneous rates of inflation expectations generated by our proposed framework and those observed in the data (i.e., Figure 1).

4.1 Standard Kalman filter

4.1.1 Setup

The economy is composed of different cohorts indexed by i. These cohorts are heterogeneous in their dates of birth and the inflation history they have experienced. Since there is no heterogeneity within cohorts, a single representative agent summarizes the situation of each one of these groups. In a given period $t+1$, the level of inflation π_{t+1} is defined according to the following random walk process[10]

[10]We opt for a random walk process instead of an AR(1) specification because the data cannot reject the hypothesis that the monthly inflation rate has a unit root. See Pivetta and Reis (2007) for a discussion on the persistence of the (quarterly) inflation rate in the US. To complete the analysis, in Appendix C, we show the model's results when the inflation series follows an AR(1) process. These results are very similar to those found under the random walk assumption.

$$\pi_{t+1} = \pi_t + \varepsilon_t,$$

where ε_t is a normally independent and identically distributed inflation shock. We assume that agents wish to forecast the future inflation rate π_{t+1}, but they only observe a noisy signal of this variable. In other words, the agents face a standard signal extraction problem. To simplify the analysis, in a given period t, we assume that the signal s_t is defined as

$$s_t = \zeta \pi_{t+1} + \upsilon_t,$$

where the coefficient $\zeta \geq 0$ denotes the pass-through existing between the unobserved variable π_{t+1} and its corresponding signal s_t, and υ_t is a signal shock. We assume that this shock is a normally independent and identically distributed variable. Moreover, to consider the existence of some elements causing movements in both the observed signal and the unobserved variable, we allow for a non-zero covariance between both shocks. More precisely, we assume

$$\begin{pmatrix} \varepsilon_t \\ \upsilon_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon\upsilon} \\ \sigma_{\varepsilon\upsilon} & \sigma_\upsilon^2 \end{pmatrix} \right).$$

As a further simplification, we assume that there is no private information in the model. In other words, all of the agents receive exactly the same signal. Since the agents face a standard signal extraction problem, we assume that they generate a forecast of the inflation variable using the corresponding conditional expected value of the variable. More precisely, given their information set in period t, the agents apply a linear Kalman filter to forecast inflation in period $t+1$. Therefore, the predicted value of the inflation variable is given by

$$\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] = \left(1 - \zeta K\right) \mathbb{E}_{i,t-1}^{KF}\left[\pi_{t+1}\right] + K s_t, \tag{1}$$

where K denotes the Kalman gain.[11] The Kalman filter approach allows us to characterize the forecasted distribution of the unobserved variable π_{t+1} in any period t conditional on agents' past and current signals $\{s_j\}_{j \in [0,t]}$. Notice that when the signal is perfectly revealing about the true state of the variable π_{t+1}, we have $\zeta = 1$, $\upsilon_t = 0$ in every period t, and $\sigma_{\varepsilon\upsilon} = 0$. As a conclusion, we obtain $K = \zeta K = 1$ and $\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] = s_t = \pi_{t+1}$. The presence of a signal noise induces $K, \zeta K \in [0, 1)$ even without a correlation between both error terms.

[11] As usual, this signal-to-noise ratio is defined such that it minimizes the variance of the prediction error associated with the unobserved variable, i.e., $\pi_{t+1} - \mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right]$. The Kalman gain that solves this optimization problem is a function of the covariance existing between the error associated with the observed signal and the unobserved variable, and the constant $\Sigma_{t+1|t-1} = \mathbb{V}\text{ar}\left[\pi_{t+1} - \mathbb{E}_{i,t-1}^{KF}\left[\pi_{t+1}\right]\right]$ (see Cheung 1993 for an example of a Kalman gain that considers these terms). Regarding the constant $\Sigma_{t+1|t-1}$, it satisfies

$$\left(\Sigma_{t+1|t-1} - \sigma_\varepsilon^2\right)\left(\zeta^2 \Sigma_{t+1|t-1} + \sigma_\upsilon^2 + 2\zeta\sigma_{\varepsilon\upsilon}\right) - \left(\sigma_\upsilon^2 \Sigma_{t+1|t-1} - \sigma_{\varepsilon\upsilon}^2\right) = 0.$$

Regarding long-run values of inflation expectations, from the Kalman-based prediction equation and using the random walk structure associated with the inflation variable, we conclude that given $h \geq 1$, we must have

$$\mathbb{E}_{i,t}^{KF}\left[\pi_{t+h}\right] = \mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right].$$

Finally, and considering $\gamma = (1 - \zeta K) \in [0, 1]$, the Kalman filter prediction can be written recursively as

$$\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] = \gamma^{t+1}\mathbb{E}_{i,-1}^{KF}\left[\pi_0\right] + K \sum_{j=0}^{t} \gamma^{t-j}\left(\zeta\pi_{j+1} + \upsilon_j\right)$$

Therefore, using this simple version of the model, we conclude that higher values of past inflation imply a higher forecasting value of this same variable. However, agents' personal experiences are not associated with heterogeneity in expectations. According to this model, in any period t, agents who lived through episodes of high inflation forecast an inflation value identical to those who lived through episodes of low inflation. Given a starting point assumption where every agent observes the initial level of inflation, i.e. $\mathbb{E}_{i,-1}^{KF}\left[\pi_0\right] = \pi_0$ for every agent i, we conclude that $\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] = \mathbb{E}_{t}^{KF}\left[\pi_{t+1}\right]$ must hold for every agent. In what follows, to simplify the analysis, we assume $\zeta = 1$.

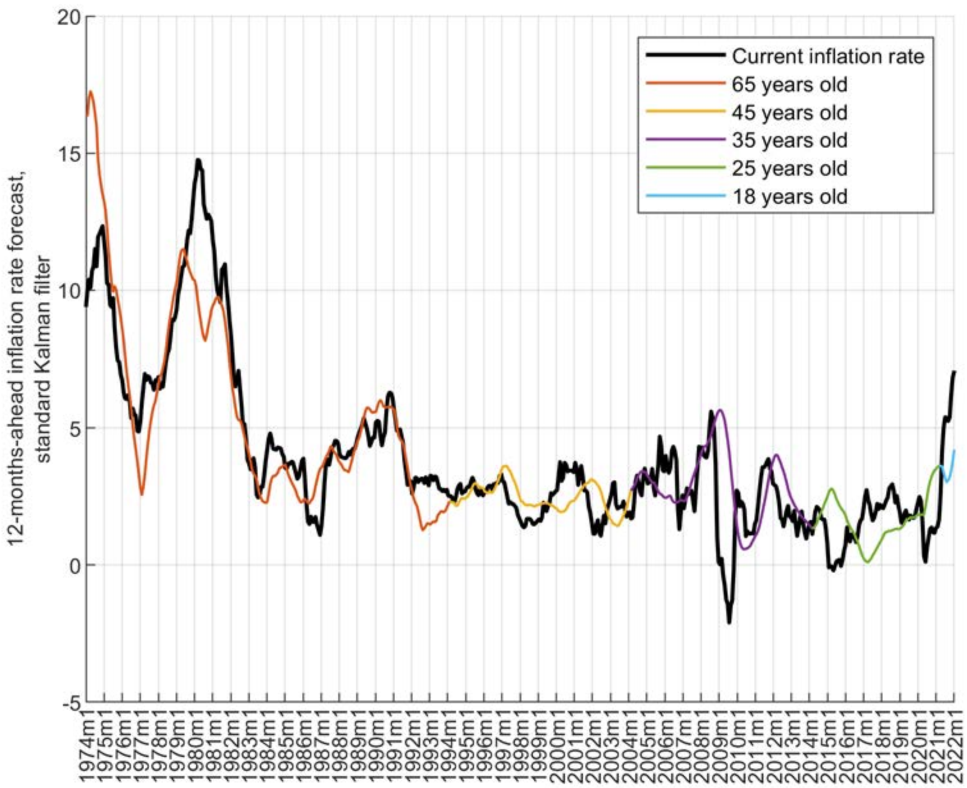
To identify an appropriate signal for the empirical counterpart of the expectations formation model, we follow the evidence presented in D'Acunto et al. (2021). This paper shows that agents use their consumption experience to form expectations. More specifically, Campos, McMain, and Pedemonte (2022), using the University of Michigan's Survey of Consumers (MSC), conclude that consumers weigh the food components of the CPI highly when forming inflation expectations. Dietrich (2022) finds similar evidence using different data sources. Therefore, we use the food component of the CPI as a shared inflation signal for consumers. More precisely, we use the previous month's food component of the CPI of the month when the agents forecast aggregate inflation. For example, if an agent forecasts aggregate inflation in December, and we presume that consumers make this prediction at the beginning of the month, we assume that this agent considers November's food inflation to make her forecast. In sum, for the empirical counterpart of the expectations formation model, we assume $s_t = \pi_{t-1}^{food}$ where π_{t-1}^{food} denotes food inflation in period $t - 1$.

We set $\sigma_{\varepsilon}^2 = 0.15$, $\sigma_{\upsilon}^2 = 4.09$ and $\sigma_{\varepsilon\upsilon} = -0.03$ from monthly data on the inflation rate. Given this calibration, we obtain $K = 0.1751$.

4.1.2 Forecasting exercise

We now perform a forecasting exercise using monthly inflation data and distinguishing agents by cohorts. Given the recursive structure of the Kalman filter, and to initialize the forecasting process of each cohort, we assume that in the period in which the cohort representative agent reaches adulthood and begins forecasting, she uses the previous period's Kalman filter expected value as a starting point. We denote the period when agent i starts forecasting as period k_i. Given the starting point assumption where the initial level of inflation is common knowledge, we have $\mathbb{E}^{KF}_{i,k_i-1}[\pi_{k_i}] = \mathbb{E}^{KF}_{k_i-1}[\pi_{k_i}]$ for every agent i. Figure 4 presents the 12-months-ahead inflation forecasts by different cohorts according to the standard Kalman filter. This figure plots the actual inflation rate and the forecast made by different selected cohorts.

Figure 4: Standard Kalman-filter-based inflation forecasts by cohort



Note: Figure shows the Kalman filter forecast for the common component for selected cohorts, differentiated by their age in 2021. We further assume each cohort starts forecasting when they become 18 years old.

As expected, the standard Kalman filter does not generate the heterogeneous pattern in inflation expectations observed in the data (i.e., Figure 1). In other words, the rate of inflation expectations evolves following an identical process across cohorts. We need to move to a more sophisticated framework to replicate the facts observed in the data.

4.2 Diagnostic Kalman filter

4.2.1 Setup

In this section, we depart from the standard Kalman filter framework by adopting the model of non-Bayesian beliefs known as diagnostic expectations. Following Bordalo, Gennaioli, and Shleifer (2018), Bordalo et al. (2019) and Bordalo et al. (2020), we denote the true conditional distribution of the unknown inflation variable in a given period t as $f\left(\pi_{t+1} \mid \mathcal{I}_{i,t}\right)$. The term $\mathcal{I}_{i,t}$ denotes the information available to agent i up to the current period t. Given this definition, we assume that the diagnostic belief distribution of inflation for agent i is given by

$$f_{i,t}^{\theta}\left(\pi_{t+1}\right) = f\left(\pi_{t+1} \mid \mathcal{I}_{i,t}\right) D_{i,t}^{\theta}\left(\pi_{t+1}\right) Z_{i,t}, \tag{2}$$

where

$$D_{i,t}^{\theta}\left(\pi_{t+1}\right) = \left[\frac{f\left(\pi_{t+1} \mid \mathcal{I}_{i,t}\right)}{f\left(\pi_{t+1} \mid \mathcal{I}_{i,t}^{ref}\right)^{t-k_i}}\right]^{\theta}.$$

In this setup, the diagnostic parameter $\theta \in \mathbb{R}$ governs the level of distortion that the likelihood ratio $\left[D_{i,t}^{\theta}\left(\pi_{t+1}\right)\right]^{\frac{1}{\theta}}$ introduces into agents' beliefs. The normalizing parameter $Z_{i,t}^{-1} = \int f_{i,t}^{\theta}\left(\pi_{t+1}\right) d\pi_{t+1}$ is a constant that ensures that the diagnostic distribution $f_{i,t}^{\theta}\left(\pi_{t+1}\right)$ integrates to one in every period t and for every agent i. In this setup, agent i compares hers current information set $\mathcal{I}_{i,t}$ against a referential information set $\mathcal{I}_{i,t}^{ref}$. Later we show that this referential information set relates to agents' i past inflation experiences. As mentioned above, the parameter θ captures the level of distortion associated with the model. Under a standard Kalman filter framework, we have $\theta = 0$, which implies $D_{i,t}^{\theta}\left(\pi_{t+1}\right) = 1$. In this case, there is no distortion in beliefs, and $f_{i,t}^{\theta}\left(\pi_{t+1}\right) = f\left(\pi_{t+1} \mid \mathcal{I}_t\right)$. When $\theta \neq 0$, beliefs are distorted.

Notice that the no private information assumption implies that the set of information associated with the true conditional distribution is equal for everyone. In other words, we have $\mathcal{I}_{i,t} = \mathcal{I}_t$ for every agent i. Therefore, under the proposed signal extraction framework, in any period t, the expected value associated with the true conditional distribution of the unknown inflation variable is common to every agent and corresponds to $\mathbb{E}_t^{KF}\left[\pi_{t+1}\right]$. Given the normality assumption on the error term of the inflation process, the true conditional distribution satisfies $f\left(\pi_{t+1} \mid \mathcal{I}_t\right) \sim \mathcal{N}\left(\mathbb{E}_t^{KF}\left[\pi_{t+1}\right], \sigma_\pi^2\right)$. As we explain below, this normality result implies that the distribution associated with the referential information set $\mathcal{I}_{i,t}^{ref}$ is normal too. Given both normality results, from Equation 2, we can show that the diagnostic distribution of agent i satisfies $f_{i,t}^{\theta}\left(\pi_{t+1}\right) \dot{\sim} \mathcal{N}\left(\mathbb{E}_{i,t}^{\theta}\left[\pi_{t+1}\right], \sigma_\pi^2\right)$, where the mean value of this distribution has the following linear structure

$$\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+1}\right] = \mathbb{E}^{KF}_{t}\left[\pi_{t+1}\right] + \theta\left(\mathbb{E}^{KF}_{t}\left[\pi_{t+1}\right] - \mathbb{E}^{ref}_{i,t}\left[\pi_{t+1}\right]\right), \tag{3}$$

where $\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+1}\right]$ is the diagnostic-distorted forecast associated with the diagnostic belief distribution $f^{\theta}_{i,t}\left(\pi_{t+1}\right)$, and $\mathbb{E}^{ref}_{i,t}\left[\pi_{t+1}\right]$ is the expected value obtained according to the distribution associated with the referential information set $\mathcal{I}^{ref}_{i,t}$. We define this linear composition of the standard Kalman filter as our diagnostic-augmented Kalman filter. When $\theta > 0$, agents over-react to the information contained in the current information set $\mathcal{I}_t$ and not in the referential information set $\mathcal{I}^{ref}_{i,t}$. Given the structure that we assume for $\mathbb{E}^{ref}_{i,t}\left[\pi_{t+1}\right]$, agents over-react to the information just received compared to their references when $\theta > 0$. If $\theta < 0$, agents under-react to the information just received, placing more weight on their references. As explained above, when $\theta = 0$, there are no distortions in beliefs, and we conclude $\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+1}\right] = \mathbb{E}^{KF}_{t}\left[\pi_{t+1}\right]$ for every cohort i.

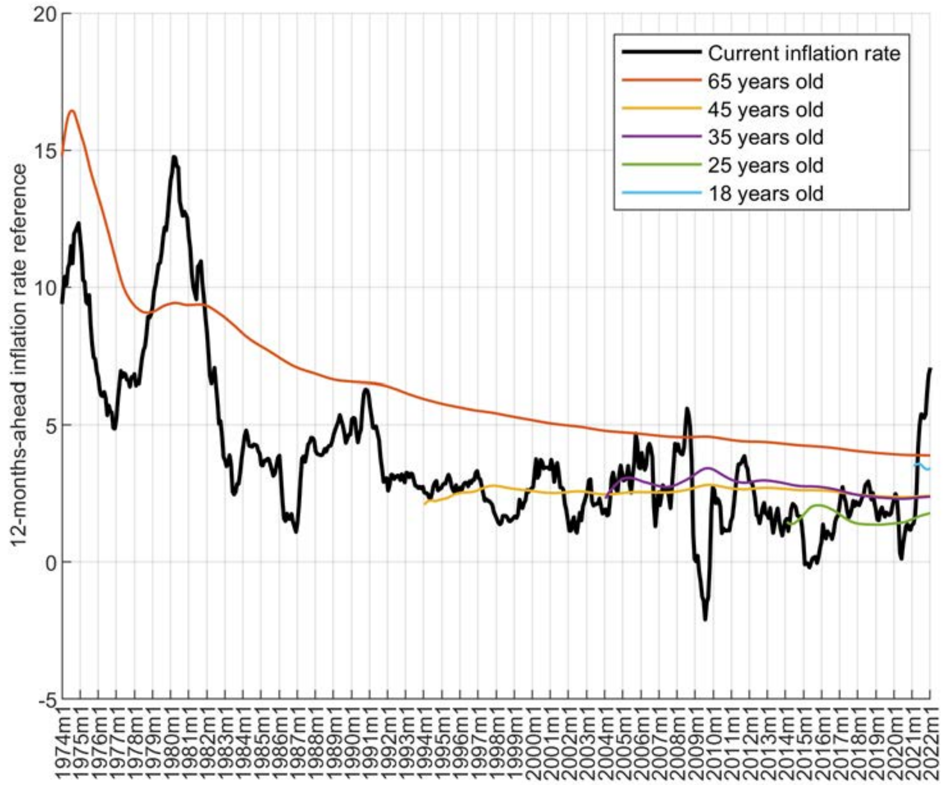
Until this point, our definitions have been history-independent. The standard Kalman filter is Markovian in the sense that it only needs the belief from the previous period, but this mechanism is not able to reproduce the empirical facts. Thus, we now introduce the role of the past through the reference term $\mathbb{E}^{ref}_{i,t}\left[\pi_{t+1}\right]$. For a representative agent of cohort i, we define the reference term as

$$\mathbb{E}^{ref}_{i,t}\left[\pi_{t+1}\right] = \frac{\sum_{j=1}^{t-k_i} \mathbb{E}^{KF}_{i,t-j}\left[\pi_{t+1}\right]}{t-k_i}, \tag{4}$$

where k_i is the period in which cohort i reaches adulthood and starts forecasting. This way, the reference term contains all the expectations agent i had in the past about the future inflation rate.

In Figure 5 we show how the inflation rate reference defined in Equation 4 evolves for different cohorts. We see that older cohorts, which have gone through episodes of higher inflation in their lifetimes, have higher reference points when compared to the younger cohorts, which have not experienced inflationary episodes.

Figure 5: Inflation rate reference by cohort



Note: Figure shows the references for selected cohorts obtained according to the Kalman filter and given the history of inflation experienced by the corresponding age group. Selected cohorts are differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

After agent i forecasts inflation for period $t+1$, the next step is to forecast its future values. Given the random walk structure associated with the inflation variable, and considering $h \geq 1$, we concluded $\mathbb{E}_t^{KF}[\pi_{t+h}] = \mathbb{E}_t^{KF}[\pi_{t+1}]$. Since the reference term is a linear composition of expected values associated with the true conditional distribution, we must have $\mathbb{E}_{i,t}^{ref}[\pi_{t+h}] = \mathbb{E}_{i,t}^{ref}[\pi_{t+1}]$. Therefore, when $h \geq 1$, we observe

$$\mathbb{E}_{i,t}^{\theta}[\pi_{t+h}] = \mathbb{E}_{i,t}^{\theta}[\pi_{t+1}] \,.$$

4.2.2 Estimation and forecasting exercise

Before performing a forecasting exercise based on the diagnostic Kalman filter, we need to know the value of the diagnostic parameter θ. In this section we propose a way of estimating this directly from the data.

We begin with the diagnostic Kalman filter from Equation 3, rewritten for the forecast agent i makes for period $t+12$, so

$$\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+12}\right] = \mathbb{E}^{KF}_{i,t}\left[\pi_{t+12}\right] + \theta\left(\mathbb{E}^{KF}_{i,t}\left[\pi_{t+12}\right] - \mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right]\right).$$

However, we know from Section 4.1 that under our current assumptions it is reasonable to assume that $\mathbb{E}^{KF}_{i,t}\left[\pi_{t+12}\right] = \mathbb{E}^{KF}_{t}\left[\pi_{t+12}\right] \ \forall i$. Then, the diagnostic Kalman filter becomes

$$\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+12}\right] = \mathbb{E}^{KF}_{t}\left[\pi_{t+12}\right] + \theta\left(\mathbb{E}^{KF}_{t}\left[\pi_{t+12}\right] - \mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right]\right).$$

Rearranging terms, we obtain

$$\mathbb{E}_{i,t}\left[\pi_{t+12}\right] = \left(1+\theta\right)\mathbb{E}^{KF}_{t}\left[\pi_{t+12}\right] - \theta\mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right]. \tag{5}$$

We now explain how to take Equation 5 to the data. First, $\left(1+\theta\right)\mathbb{E}^{KF}_{t}\left[\pi_{t}\right]$ is common across all cohorts, so it can be captured by a time fixed effect γ_t. Second, for the distorted inflation expectation $\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+12}\right]$ we use the 12-months-ahead forecasts for each agent m from cohort i from the SCE $\mathbb{E}^{SCE}_{m,i,t}\left[\pi_{t+12}\right]$. Lastly, for $\mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right] = \frac{\sum_{j=1}^{k_i}\mathbb{E}^{KF}_{i,t-j}\left[\pi_{t+12}\right]}{t-k_i}$ we go back to the standard Kalman filter from Section 4.1 and recover the terms $\mathbb{E}^{KF}_{i,t-j}\left[\pi_{t+12}\right]$, which are the optimal forecasts under the given setup.

With this, we regress

$$\mathbb{E}^{SCE}_{m,i,t}\left[\pi_{t+12}\right] = \gamma_t + \varphi\mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right] + \varepsilon_{m,i,t}. \tag{6}$$

We present the results in Table 2. Column 1 shows the main specification, from which we obtain $\theta = -\widehat{\varphi} = -0.324$. Because the diagnostic parameter $\theta < 0$, this means agents under-react to the current information with respect to their reference information sets. In other words, when agents make their forecasts, they put more weight on their reference sets (i.e. their past history, their priors) than on the news they receive in the current period. In the other columns of Table 2 we show that the under-reaction result is robust to the inclusion of control variables in the regression.

Table 2: Diagnostic parameter estimation

	(1)	(2)	(3)	(4)
$\mathbb{E}_{i,t}^{ref}\left[\pi_{t+12}\right]$	0.324*** (0.030)	0.362*** (0.027)	0.266*** (0.029)	0.231*** (0.026)
Time FE	Yes	Yes	Yes	Yes
Controls	No	Gender, commuting zone	Gender, commuting zone, HH income	Gender, commuting zone, HH income, educational degree
Observations	101,256	101,239	101,239	101,239
R-squared	0.092	0.149	0.171	0.185

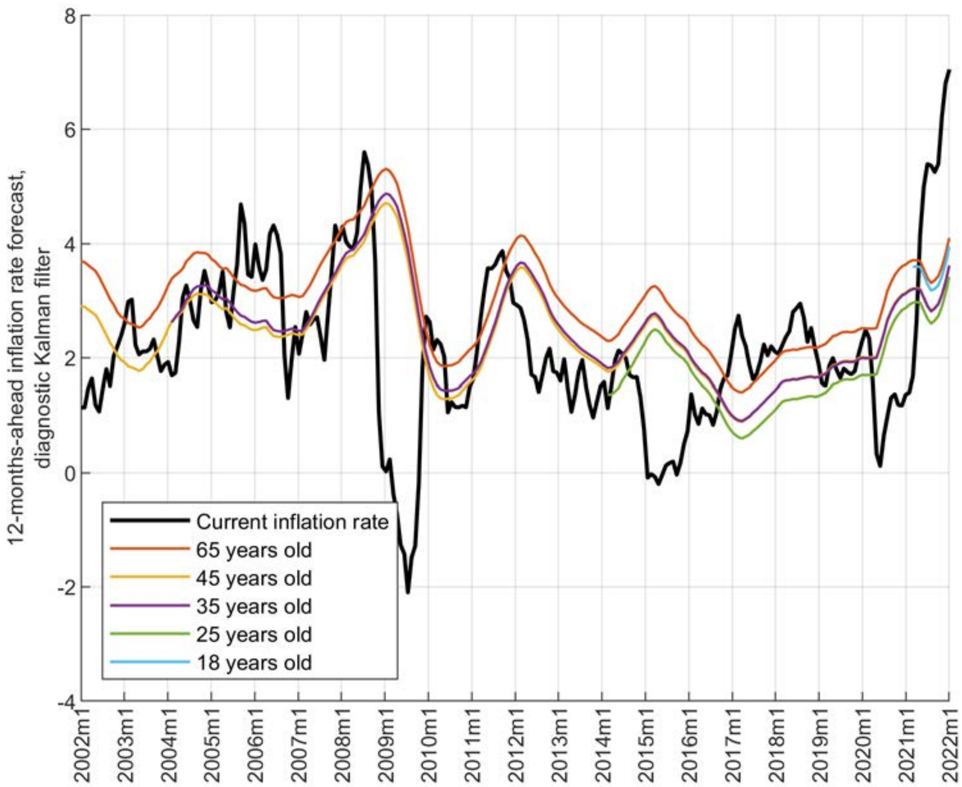
Note: Table shows results of Regression (6). $\mathbb{E}_{i,t}^{ref}\left[\pi_{t+12}\right]$ is the reference constructed for a respondent of age i as explained in the main text. Column (1) has only a time fixed effect as an additional control. Columns (2), (3) and (4) add different levels of controls. Robust standard errors in parentheses. Standard errors clustered by age. Dependent variable trimmed at 10 percent and 90 percent in each period. *** p<0.01, ** p<0.05, * p<0.1.

This result contrasts with the diagnostic expectations literature (for instance, Bordalo et al. 2020), where a positive θ is the usual result. However, most of the previous empirical literature on diagnostic expectations relies on surveys of professional forecasters, who have a better knowledge and are better informed about the economy than the households surveyed in the SCE. Thus, a conclusion would be that professional forecasters do over-react to current news, while households under-react to news because they rely more on their history.

With this estimate we perform a forecasting exercise using the diagnostic Kalman filter. Figure 6 shows the 12-months-ahead inflation forecasts for different cohorts. We can see that the diagnostic Kalman filter is able to generate a heterogeneous pattern across cohorts, similar to the one we saw in the data in Figure 1.[12] First, older cohorts show higher inflation expectations than the rest of the cohorts throughout most of the sample, based on their experiences of high inflation in the 60s, 70s, and early 80s. Second, the intermediate cohorts show low inflation expectations when compared to the other cohorts, because they went through the stable and low inflation rates of the 90s, 00s, and 10s. The youngest cohort shows the highest inflation expectations after being exposed to the high inflation rates of 2021.

[12] We show the diagnostic forecasts for the full sample in Figure A.1 in Appendix A.

Figure 6: Diagnostic Kalman-filter-based inflation forecasts by cohort



Note: Figure shows forecasts for selected cohorts according to the Kalman-filter-augmented expectations and considering the estimate for θ from Column 1 of Table 2. Selected cohorts differentiated by their age in 2021. We further assume each cohort starts forecasting when they become 18 years old.

As a way of checking the external validity of our results, in Appendix D we perform the same exercises but using data from the Consumer Expectations Survey of the European Central Bank, which contains observations for six countries: Belgium, France, Germany, Italy, the Netherlands, and Spain. We find evidence that supports our main findings. We find that inflation expectations are also heterogeneous in Europe and are explained partially by the history of consumers' inflation experiences. In addition, we also find support for the use of a diagnostic Kalman filter as a way of modeling heterogeneous inflation expectations.

Moreover, thanks to the cross-country panel structure, we are able to control for a common cohort fixed effect, as in Hajdini et al. (2022a). This is important, as common cohort characteristics (i.e. different patterns of inflation exposure over the life cycle) could affect our results. By controlling for cohort fixed effect we rule out those common cohort characteristics and exploit differences in the inflation experienced by the different cohorts across different countries. We find similar results after adding these controls, implying that the heterogeneity does not stem from common cohort characteristics, but from different inflation experiences in different countries. Thus, we conclude that our findings are also valid for Europe.

4.2.3 Goodness of fit

In this subsection we show how our proposed diagnostic Kalman filter expectations, which combine time variation and individual variation, compare to the data.

Column 1 of Table 3 shows that the slope between the survey forecast and the forecast produced by the diagnostic Kalman filter is 0.899. This confirms that our diagnostic formulation for inflation expectations is effective in forecasting consumers' inflation expectations. Our formulation has two components: one that is coming from the Kalman filter, with a signal coming from common food inflation data, and the second coming from the past references of cohorts and a coefficient that is estimated using the data. In that sense, the time variation from our inflation expectations measure is not being informed by the individual data, as we use a time fixed effect for the coefficient estimation. After considering the time and cross-section variation, our estimate is able to provide a good prediction of heterogeneous inflation expectations.

An alternative measure to explain the heterogeneous inflation expectations from the data is the lifetime average inflation rate by cohort $\overline{\pi_{i,t}} = \frac{\sum_{j=0}^{k_i} \pi_{t-j}}{t-k_i+1}$. Column 2 of Table 3 shows that the history of inflation by cohort, by itself, is also able to predict part of the variation in the data.

We can make the diagnostic Kalman filter compete with the lifetime average inflation rate to see which measure better predicts the forecasts we see in the data by estimating the following regression:

$$\mathbb{E}_{m,i,t}^{SCE}\left[\pi_{t+12}\right] = \omega_1 \mathbb{E}_{i,t}^{\theta}\left[\pi_{t+12}\right] + \omega_2 \overline{\pi_{i,t}} + \varepsilon_{i,t}. \tag{7}$$

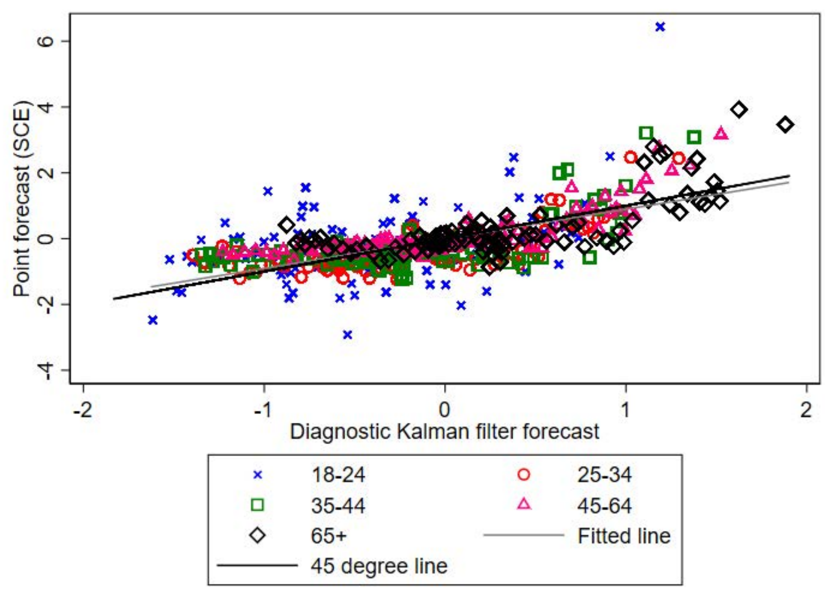
Column 3 of Table 3 shows that in a horserace our diagnostic forecast is superior to the lifetime average inflation rate for explaining the observed heterogeneous inflation expectations we find in the data. The coefficient for our diagnostic measure is close to one and statistically significant, while the coefficient for the history of inflation by cohort goes close to zero and becomes statistically insignificant.

Table 3: Goodness of fit

	(1)	(2)	(3)
$\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+12}\right]$	0.899***		0.888***
	(0.037)		(0.039)
$\overline{\pi_{i,t}}$		0.230***	0.027
		(0.028)	(0.030)
Observations	101,256	101,256	101,256
R-squared	0.036	0.004	0.036

Note: Table shows results of Regression (7). The dependent variable is consumers' inflation expectations according to the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York. $\mathbb{E}^{\theta}_{i,t}\left[\pi_{t+12}\right]$ is our estimated measure of inflation expectations. $\overline{\pi_{i,t}}$ is average inflation expectations. Standard errors clustered at the date-of-birth level in parentheses. Dependent variable trimmed at 10 percent and 90 percent in each period. *** p<0.01, ** p<0.05, * p<0.1.

Figure 7: Observed inflation forecasts and diagnostic Kalman filter forecasts



Note: Figure shows binned scatterplot across diagnostic Kalman filter forecasts (x-axis) and point forecast inflation expectations according to the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York (y-axis). Variables demeaned by the intercept. Data go from June 2013 to December 2021. SCE variable trimmed at 10 percent and 90 percent in each period.

Figure 7 visually presents the results of Column 1 in Table 3. We can see that the slope between the regression and a 45-degree line are very close. Our diagnostic measure can effectively model the time and cross sectional variation of consumers' inflation expectations.

Overall, we show that our diagnostic measure shows a very good fit with the data and that we are able to replicate heterogeneous inflation expectations at the individual level, a complicated object, with a relatively simple model of expectations. More importantly, as it is data driven, we can easily incorporate this diagnostic measure into a general equilibrium model to see the implications of this expectations formation process. In the next section we propose a way to include these findings, using our measure, in a DSGE model to study the aggregate implications of heterogeneity in consumers' expectations.

5 Aggregate implications of heterogeneous expectations

Our goal in this section is to present an overlapping generations monetary model that replicates the heterogeneity in the observed inflation expectations (i.e., Figure 1). To do so, we assume that agents follow the diagnostic Kalman filter introduced in Section 4.2 when forecasting future variables as similarly proposed in Bianchi, Ilut, and Saijo (2021) and L'Huillier, Singh, and Yoo (2021). The long memory inherent in this approach allows for different past experiences to shape different inflation expectations across cohorts.

5.1 Households

On the demand side, we assume that the economy is populated by an infinite number of cohorts. Every cohort is composed of a continuum of households, all of which can be summarized by a representative agent. The cohorts are heterogeneous in their age and past inflation experiences. For modeling the different cohorts, we follow the perpertual youth approach of Blanchard (1985) and Yaari (1965). This means that households are uncertain about the date on which they will die. All they know is that they face a rate of mortality λ every period. At the same time, every period a new cohort of size λ is born. Therefore, in a given period t, the size of a cohort born in period k is $\lambda\left(1-\lambda\right)^{t-k}$.

All households are forward looking, so they wish to forecast future values for the output gap and the inflation rate. We assume that households form their expectations using the diagnostic Kalman filter from Equation 3. Therefore, their expectations will be influenced by their past experiences. This also means that all the assumptions from Section 4.2 apply here. First, agents do not fully understand the model that governs the economy, so they assume that both the output gap and the inflation rate behave as a random walk. Second, agents cannot directly observe either the current output gap or the current inflation rate, but they instead receive a signal. With this, agents form diagnostic forecasts about the output gap and the inflation rate.

5.1.1 Consumption basket

The representative household from cohort i consumes a consumption basket $C_{i,t}$, composed of a continuum of $C_{i,t}\left(j\right)$ goods indexed by j. This basket is defined as

$$C_{i,t} = \left(\int_0^1 C_{i,t}\left(j\right)^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}},$$

where ε is the elasticity of substitution in the CES basket.

Maximizing this basket subject to a standard budget constraint gives the following first-order condition:

$$C_{i,t}\left(j\right) = \left(\frac{P_t\left(j\right)}{P_t} \right)^{-\varepsilon} C_{i,t},$$

which represents the demand for good j and where $P_t\left(j\right)$ is the price of the good and P_t is the price level.

5.1.2 Utility maximization

A representative household of cohort i solves

$$\max \left[\frac{C_{i,t}^{1-\sigma}}{1-\sigma} - \frac{L_{i,t}^{1+\eta}}{1+\eta} \right] + \sum_{j=1}^{\infty} \beta^{j-t} \left(1-\lambda\right)^{j-t} \mathbb{E}_{i,t}^{\theta} \left[\frac{C_{i,t+j}^{1-\sigma}}{1-\sigma} - \frac{L_{i,t+j}^{1+\eta}}{1+\eta} \right],$$

subject to

$$P_t C_{i,t} + \left(1-\lambda\right) \frac{B_{i,t+1}}{\left(1+i_t\right)} = W_t L_{i,t} + B_{i,t} + T_{i,t},$$

where $C_{i,t}$ is consumption, $L_{i,t}$ is the labor supply, $B_{i,t}$ are nominal savings, P_t is the price level, W_t are the nominal wages, $T_{i,t}$ are transfers, and i_t is the nominal interest rate. Also, β is the discount factor, σ is the intertemporal elasticity of substitution, and η is the inverse of the Frisch elasticity.

The transfers $T_{i,t}$ are crucial to our model, as they incorporate two different mechanisms. First, as in Blanchard (1985) and Yaari (1965) we assume that households insure themselves to receive a flow of income every period they are alive. Then, when they die, the insurance company takes away any wealth residual. Thus, we do not have to worry about accidental bequests. Second, as in Mankiw and Reis (2006), we assume that the flow of income households receive each period from the insurance company is such that households start each period with the same wealth and that the nominal savings market clears. Therefore, we do not have to worry about the wealth distribution. Lastly, as a way of closing the model, the transfers also incorporate the benefits coming from firms that produce intermediate goods.

As in Bianchi, Ilut, and Saijo (2021) and L'Huillier, Singh, and Yoo (2021) we introduce diagnostic expectations in a general equilibrium setting. The expectations operator households use is $\mathbb{E}^{\theta}_{i,t}[.]$, which works under the assumptions of Section 4.2.[13]

We also assume that for any current variable X_t we have that $\mathbb{E}^{\theta}_{i,t}[X_t] = X_t$ and that for any lagged variable X_{t-h} we have that $\mathbb{E}^{\theta}_{i,t}[X_{t-h}] = X_{t-h}$.

The first-order conditions are

$$\frac{1}{C_{i,t}} = \beta (1+i_t) \mathbb{E}^{\theta}_{i,t}\left[\frac{1}{(1+\pi_{t+1}) C_{i,t+1}}\right], \tag{8}$$

$$L^{\eta}_{i,t} = \frac{W_t}{P_t C_{i,t}}, \tag{9}$$

where the first equation is the Euler equation and the second one denotes the labor supply.

Additionally, the transversality condition is

$$\lim_{T\to\infty} \frac{(1-\lambda)^T}{\Pi_{h=0}^{T-1}(1+i_{t+h})} B_{i,t+T} = 0.$$

5.1.3 Incorporating the diagnostic Kalman filter

The log-linearization of Equation 8, following the diagnostic Kalman filter from Equation 3, gives

$$c_{i,t} = \begin{array}{c} \left\{\mathbb{E}^{KF}_t[c_{i,t+1}] - \frac{1}{\sigma}\left(i_t - \mathbb{E}^{KF}_{i,t}[\pi_{t+1}]\right)\right\} \\ +\theta\left\{\left(\mathbb{E}^{KF}_t[c_{i,t+1}] - \mathbb{E}^{ref}_{i,t}[c_{i,t+1}]\right) + \dfrac{1}{\sigma}\left(\mathbb{E}^{KF}_t[\pi_{t+1}] - \mathbb{E}^{ref}_{i,t}[\pi_{t+1}]\right)\right\} \end{array}, \tag{10}$$

where the lowercases denote deviations from the steady state.[14] This is the IS curve for a given cohort i.

[13] We assume that the assumptions for the diagnostic expectations operator also work with the output gap in the linearized model. Thus, agents believe the output gap behaves as a random walk process and forecast the future output gap using the diagnostic Kalman filter. Moreover, while in Section 4.1 we used the lag of the inflation rate of the food component of the CPI as a signal, here we use the lag of the inflation rate. This is because we only have one final consumption good in the model.

[14] An intermediate step in the log-linearization of Equation 8 results in

5.1.4 Aggregation

In our log-linearized economy, the aggregate consumption gap c_t is defined as the weighted sum of all the cohort-level consumption gaps, so

$$c_t = \lambda \sum_{k=0}^{\infty} (1-\lambda)^k c_{k,t}. \tag{11}$$

Incorporating Equation 10 into Equation 11 we find

$$c_t = \begin{array}{c} \left\{-\frac{i_t}{\sigma} + \mathbb{E}_t^{KF}[c_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} + \theta\left\{\mathbb{E}_t^{KF}[c_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} \\ -\theta\lambda\sum_{k=0}^{\infty}(1-\lambda)^k\left\{\mathbb{E}_{k,t}^{ref}[c_{k,t+1}] + \mathbb{E}_{k,t}^{ref}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\}. \end{array}$$

Here, we further assume that household k, when forecasting its future individual consumption gap, believes that all the other households will behave in a similar way such that $\mathbb{E}_{k,t}^{ref}[c_{k,t+1}] = \mathbb{E}_{k,t}^{ref}[c_{t+1}]$ and

$$c_t = \begin{array}{c} \left\{-\frac{i_t}{\sigma} + \mathbb{E}_t^{KF}[c_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} + \theta\left\{\mathbb{E}_t^{KF}[c_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} \\ -\theta\lambda\sum_{k=0}^{\infty}(1-\lambda)^k\left\{\mathbb{E}_{k,t}^{ref}[c_{t+1}] + \mathbb{E}_{k,t}^{ref}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\}. \end{array}$$

Further assuming that in equilibrium the output gap $y_t = c_t$, then

$$y_t = \begin{array}{c} \left\{-\frac{i_t}{\sigma} + \mathbb{E}_t^{KF}[y_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} + \theta\left\{\mathbb{E}_t^{KF}[y_{t+1}] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} \\ -\theta\lambda\sum_{k=0}^{\infty}(1-\lambda)^k\left\{\mathbb{E}_{k,t}^{ref}[y_{t+1}] + \mathbb{E}_{k,t}^{ref}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\}. \end{array} \tag{12}$$

Equation 12 is the diagnostic IS curve in our model. It is equal to the standard IS curve plus two distortion terms. In this version of the IS curve, the past matters in the sense that current realizations are affected by the memory of the cohorts.

$$c_{i,t} = \begin{array}{c} \left\{\mathbb{E}_t^{KF}[c_{i,t+1}] - \frac{1}{\sigma}\left(i_t - \mathbb{E}_{i,t}^{KF}[\pi_{t+1}]\right)\right\} \\ +\theta\left\{\left(\mathbb{E}_t^{KF}[c_{i,t+1}] - \mathbb{E}_{i,t}^{ref}[c_{i,t+1}]\right) + \frac{1}{\sigma}\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{i,t}^{ref}[\pi_{t+1}]\right)\right\} \\ +\frac{\theta}{\sigma}\sum_{j=0}^{k_i+1}\left(\mathbb{E}_{i,t}^{\theta}[\pi_{t-j}] - \pi_{t-j}\right), \end{array}$$

where the last term results from the fact that $\mathbb{E}_{i,t}^{\theta}[X_{t+1}Z_t] \neq \mathbb{E}_{i,t}^{\theta}[X_{t+1}]Z_t$ (see, for instance, L'Huillier, Singh, and Yoo 2021). Because $\mathbb{E}_{i,t}^{\theta}[\pi_{t-j}] = \pi_{t-j}$ we drop the last term and obtain Equation 10.

5.2 Firms

On the supply side, there is a final goods producer that operates in a perfectly competitive market, which produces using a continuum of intermediate goods as inputs. There is a continuum of intermediate goods producers, each operating under monopolistic competition. These intermediate goods producers are subject to Calvo pricing frictions.

We assume that these firms follow rational expectations when setting their prices, in the sense that they are model consistent. Thus, we follow the usual derivations for firms in a New Keynesian setting, such that we obtain the standard New Keynesian Phillips curve.[15]

5.3 Monetary policy

The central bank sets the interest rate following a standard Taylor rule. Then, we have

$$\frac{(1+i_t)}{(1+\bar{i})} = \left[\frac{(1+\pi_t)}{(1+\overline{\pi})}\right]^{\chi_\pi} \left[\frac{Y_t}{\overline{Y}}\right]^{\chi_y}, \tag{13}$$

where the bars denote steady state values and χ_π and χ_y represent the central bank's reaction to deviations from the steady state of the inflation rate and output, respectively.

5.4 Summary

After log-linearizing, the model is summarized by

$$y_t = \begin{array}{c} \left\{-\frac{i_t}{\sigma} + \mathbb{E}_t^{KF}\left[y_{t+1}\right] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} + \theta\left\{\mathbb{E}_t^{KF}\left[y_{t+1}\right] + \mathbb{E}_t^{KF}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} \\ -\theta\lambda\sum_{k=0}^{\infty}(1-\lambda)^k\left\{\mathbb{E}_{k,t}^{ref}\left[y_{t+1}\right] + \mathbb{E}_{k,t}^{ref}\left[\frac{\pi_{t+1}}{\sigma}\right]\right\} + u_t^{taste}, \end{array} \tag{14}$$

$$\pi_t = \frac{(1-\phi)(1-\phi\beta)}{\phi}(\sigma+\eta)\, y_t + \beta\mathbb{E}_t\left[\pi_{t+1}\right] + u_t^{cost}, \tag{15}$$

$$i_t = \chi_\pi \pi_t + \chi_y y_t, \tag{16}$$

where Equation 14 is the diagnostic dynamic IS curve augmented with a taste shock, Equation 15 is the Phillips curve, and Equation 16 is the Taylor rule. Notice the Phillips curve follows the rational expectations operator $E_t[.]$, while the diagnostic IS curve results from following the diagnostic Kalman filter operator $E_t^{\theta}[.]$.

[15] We present the derivations in Appendix E.

We consider a cost shock u_t^{cost} and a taste shock u_t^{taste} that behave as an AR(1) process, such that

$$u_t^{cost} = \rho_{cost} u_{t-1}^{cost} + \epsilon_t^{cost}, \tag{17}$$

$$u_t^{taste} = \rho_{taste} u_{t-1}^{taste} + \epsilon_t^{taste}, \tag{18}$$

where ρ_{cost} and ρ_{taste} are the persistence parameter and ϵ_t^{cost} and ϵ_t^{taste} are the unexpected innovations.

5.5 Calibration

The model is calibrated to a monthly frequency. The parameters from Table 4 show a fairly standard calibration.

Table 4: Model calibration

Parameter	Value	Parameter	Value
β	0.9967	χ_y	0.125
η	1	ρ^{cost}	0.9
ϕ	0.9167	ρ^{taste}	0.9
σ	1	λ	0.001
ε	9	K	0.1751
χ_π	1.5	θ	-0.324

Note: Table shows the parameters used for the model. We follow a standard monthly calibration.

We calibrate the price stickiness parameter ϕ so that the expected duration of a given price quote is 12 months. We also calibrate the mortality rate λ so that the expected life span is 80 years.[16]

Regarding the diagnostic Kalman filter, we need to calibrate the Kalman gain K and the diagnostic parameter θ. We calibrate both according to the results from Sections 4.1 and 4.2. We must make an additional assumption around these two parameters. While we only used inflation rate data in the previous sections, here we assume that these parameters also hold true for the output gap.

[16]Because we assume that agents become economically active and relevant at age 18, this means agents expect to consume and work for 62 years

5.6 Simulations

5.6.1 Cost shock

Figure 8 presents the impulse response functions to a cost shock. We compare three different cases: (i) households form their expectations according to full information rational expectations (FIRE), (ii) households form their expectations according to the diagnostic Kalman filter but with no heterogeneity across cohorts,[17] and (iii) households form their expectations according to the diagnostic Kalman filter from Equation 3 where there is heterogeneity across cohorts.

First, we compare the two diagnostic cases (blue and green lines) to the FIRE case (red lines). We see that upon the shock, the output gap, determined by the diagnostic-influenced households, under-reacts when there are diagnostic expectations relative to the FIRE case. This is because diagnosticity anchors the future output gap expectations to the reference term, which is itself based on the steady state. Thus, with diagnosticity the IS curve becomes more inelastic to the shock.

With a more inelastic IS curve, rational firms are able to raise prices by more than they would under FIRE. Thus, in the diagnostic economy, the central bank must raise the interest rate more strongly than in the rational economy. When we compare the average expectations, the diagnosticity generates a hump shape with respect to the FIRE case. Moreover, the effects of the shock on expectations last longer than under FIRE.

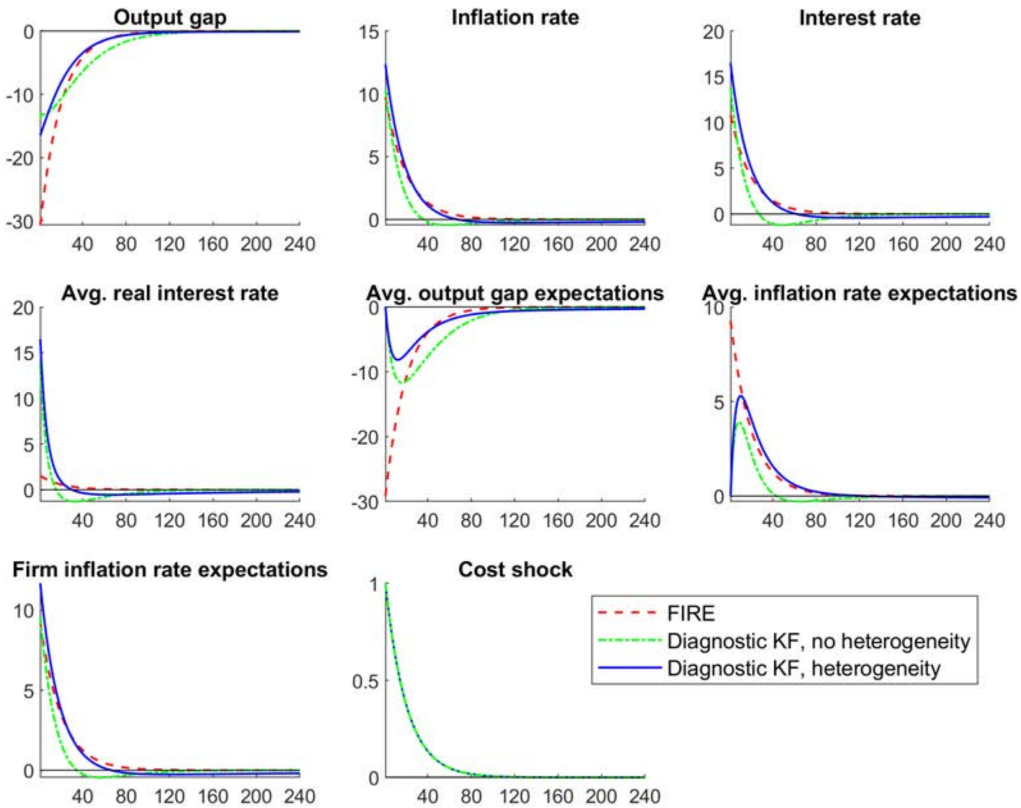
Now we turn to comparing the two diagnostic cases: the diagnostic Kalman filter with heterogeneity (blue lines) to the diagnostic Kalman filter with no heterogeneity (green lines). The under-reaction in the output gap is more notable under our diagnostic Kalman filter than with the standard diagnostic Kalman filter specification. In the former, agents are pegged to the steady state for longer and are slower to respond to the shock. When it comes to average expectations, the shock effects are lower but last much longer under our diagnostic Kalman filter. We also see that heterogeneity increases the duration of the effects of the shocks, as people have longer memories.

[17] As is standard in most of the diagnostic expectations literature (see for example Bianchi, Ilut, and Saijo 2021; L'Huillier, Singh, and Yoo 2021), the diagnostic expectations operator with no heterogeneity for a certain variable X_{t+h} would be

$$E_t^{\theta,nh}\left[X_{t+h}\right] = E_t^{KF}\left[X_{t+h}\right] + \theta\left(E_t^{KF}\left[X_{t+h}\right] - E_{t-3}^{KF}\left[X_{t+h}\right]\right),$$

where the reference term is fixed and nh stands for no heterogeneity. Here, heterogeneity does not play a role as this diagnostic expectations operator is valid for all cohorts i.

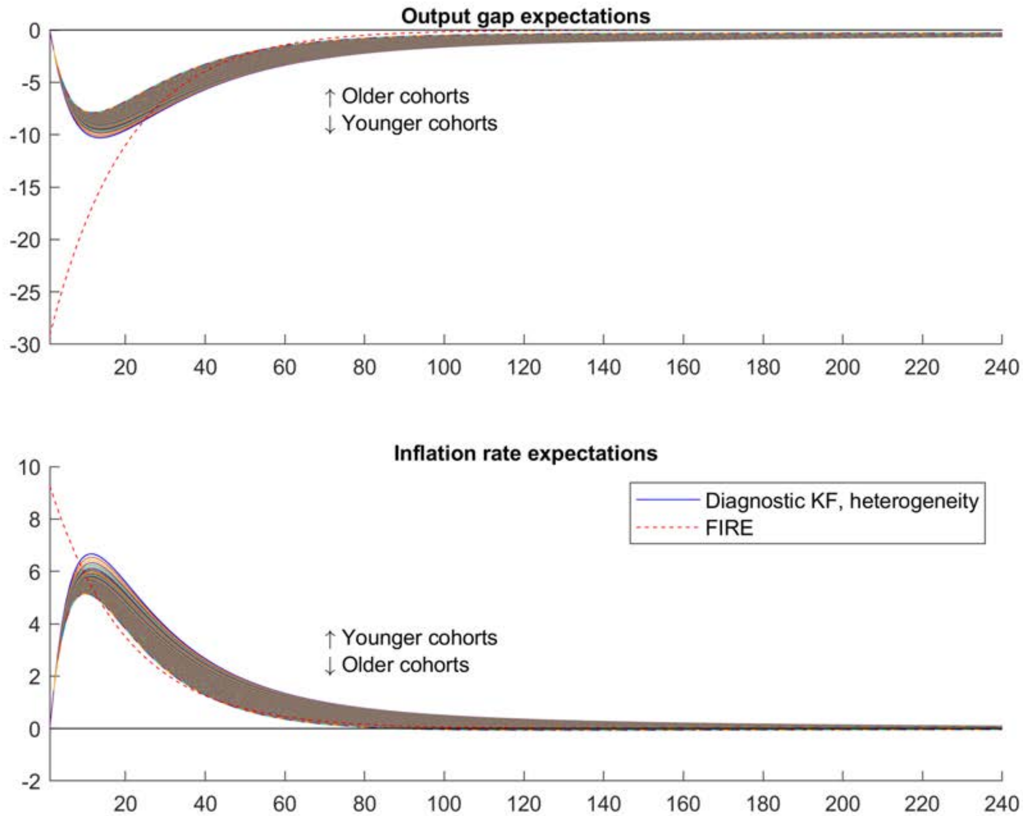
Figure 8: Impulse response functions, cost shock



Note: Figure shows impulse response functions for a selected group of variables after a cost shock. The red dashed line shows the results for the case of the full information rational expectations model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume that agents use the expectations operator $\mathbb{E}_t^{nh}[\pi_{t+1}] = \mathbb{E}_t^{KF}[\pi_{t+1}] + \theta\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{t-3}^{KF}[\pi_{t+1}]\right)$. Horizontal axis denotes months after the shock.

In Figure 9 we present the heterogeneous expectations for all cohorts under the diagnostic expectations operator. Older cohorts have their references pegged to the zero steady state, which makes them under-react to the shock when compared to the younger cohorts. On the flip side, the expectations from younger cohorts react more strongly as they are constructing their reference points with the current data.

Figure 9: Impulse response functions, inflation rate diagnostic expectations by cohort, cost shock



Note: Figure shows the heterogeneous expectations generated by the diagnostic Kalman filter. Cohorts denote age at the time of the shock. The solid lines represent different cohorts in the diagnostic Kalman filter model. The dashed red line is the full information rational expectations model. Horizontal axis denotes months after the shock.

In this case extrapolation works in a different way compared to the over-reaction model as in L'Huillier, Singh, and Yoo (2021). After a negative supply shock and coming from the steady state, older consumers are over-optimistic, as their expectations are anchored to the steady state. This means that their output expectations remain anchored, as well as their inflation expectations. This anchoring reduces the effects of the shock in terms of output. However, as firms are rational, they translate this over-optimism into higher prices, further dampening the effect of the shock on the output gap. In this case, because the shock moves the output gap and inflation in different directions, there is an attenuation effect, while inflation expectations show over-persistence, but at relatively low levels.

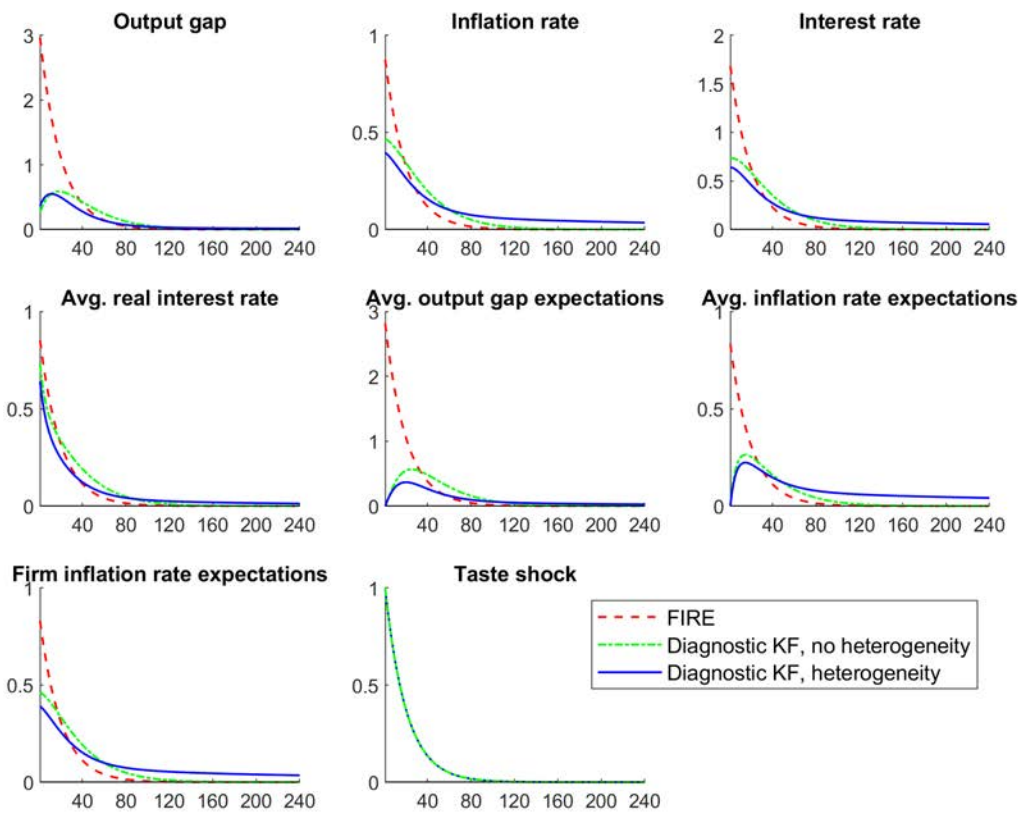
5.6.2 Taste shock

We now analyze the effects of a taste shock. Figure 10 shows the impulse response functions. We first see that the output gap under-reacts under diagnostic expectations (blue and green lines) with respect to the FIRE case (red lines). Because of this, prices under-react too when compared to the FIRE case. Therefore, the interest rate rises by less with diagnostic expectations than with FIRE.

While comparing the cases with the diagnostic Kalman filter with heterogeneity (blue lines) and the Kalman filter with no heterogeneity (green lines) we see that heterogeneity has a larger degree of anchoring to the past, but makes the effects of the shock last for longer.

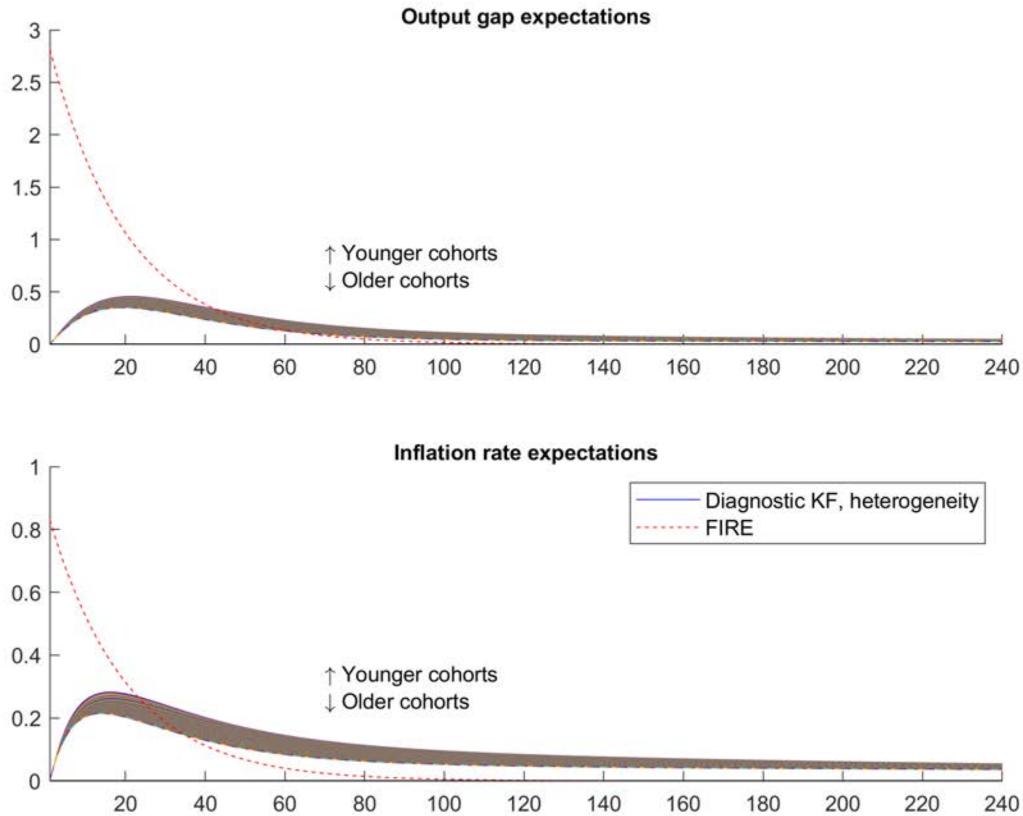
Figure 11 presents the heterogeneity in expectations across cohorts under the diagnostic Kalman filter. When it comes to the heterogeneous expectations across cohorts under the diagnostic Kalman filter, the older cohorts under-react to the shock in their expectations. Younger cohorts keep reacting more strongly, as they are building their reference on the current observations affected by the shock.

Figure 10: Impulse response functions, taste shock



Note: Figure shows impulse response functions for a selected group of variables after a taste shock. The red dashed line shows the results for the case of the full information rational expectations model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume that agents use the expectations operator $\mathbb{E}_t^{nh}[\pi_{t+1}] = \mathbb{E}_t^{KF}[\pi_{t+1}] + \theta\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{t-3}^{KF}[\pi_{t+1}]\right)$. Horizontal axis denotes months after the shock.

Figure 11: Impulse response functions, inflation rate diagnostic expectations by cohort, taste shock



Note: Figure shows the heterogeneous expectations generated by the diagnostic Kalman filter. Cohorts denote age at the time of the shock.The solid lines represent different cohorts in the diagnostic Kalman filter model. The dashed red line is the full information rational expectations model. Horizontal axis denotes months after the shock.

We find the opposite effect compared to the supply shock. In this case the shock increases output and inflation. So, consumers are over-pessimistic relative to the output gap. This reduces the pace of price adjustment, creating a smaller, but very persistent effect on inflation. Output only mildly adjusts, reducing its volatility. Given the cohort structure, inflation remains above the steady state for a prolonged period of time.

6 Optimal Taylor rules

In this section we analyze the use of an optimal Taylor rule in each of the different cases: (i) rational expectations, (ii) diagnostic expectations with no heterogeneity and (iii) diagnostic expectations with heterogeneity. The Taylor rule we use in this section is

$$i_t = \chi_\pi^* \pi_t + \chi_y^* y_t. \tag{19}$$

We assume that the central bank chooses the time-invariant parameters χ_π^* and χ_y^* such that it solves the problem

$$\min \mathbb{E}_t \left[\pi_t^2 + \vartheta y_t^2\right],$$

subject to the equations of the model in Section 5 and ϑ is the weight of the output gap in the objective function.[18] That is, the central bank, given the model setup, seeks to minimize the volatility of both the inflation rate and the output gap. Notice that we assume that the central bank has rational expectations.

The optimal parameters are dependent on which shocks exist in the model (cost or taste). Therefore, we will have two sets of parameters, one for each shock.

6.1 Optimal response to a cost shock

Figure 12 shows the impulse-response functions with the optimal parameters.[19] When responding to this shock, the central bank faces the usual trade-off between the output gap and the inflation rate. After the cost shock, the inflation rate goes up and the output gap goes down after the interest rate hikes. In this particular exercise we will have that, given the relative importance of the output gap in the objective function, the central bank will favor reducing the volatility of the inflation rate.

After the shock hits, the central bank becomes more active in both of the diagnostic cases when compared to the FIRE case. The reason behind this behavior is that memory plays a role when there are diagnostic expectations. The central bank knows that people will remember the current shock far into the future, affecting future inflation expectations. Therefore, the central bank finds it optimal to supress the inflation rate sooner rather than later.

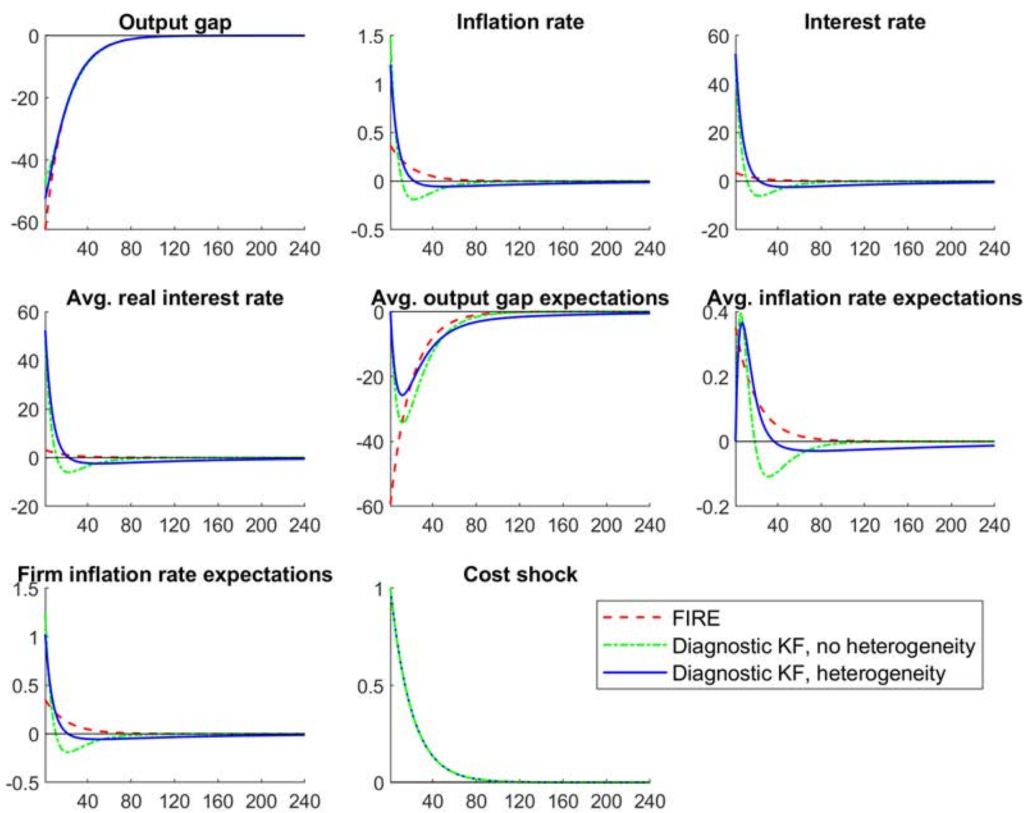
Moreover, in between the two diagnostic cases, the central bank will be the most active when there is heterogeneity. The results follow from the fact that agents have short memories when there is no heterogeneity, while they have long memories when there is heterogeneity. Thus, the central bank finds it optimal to be more active the further into the fugure the agents will remember the shock.

By being more active when there are diagnostic expectations, we see the central bank can very quickly lower the inflation expectations. While in the baseline results from Figure 8 the inflation expectations remained high for a long period, the optimal Taylor rule brings them down and even generates deflation expectations that later spill over to the observed inflation rate.

[18] Following Gali (2015) we define $\vartheta = \frac{(1-\phi)(1-\phi\beta)(\sigma+\eta)}{\phi\varepsilon} = 0.0017$.

[19] Table A.1 in Appendix A shows the optimal parameters.

Figure 12: Impulse response functions, optimal Taylor rule, cost shock



Note: Figure shows impulse response functions for a selected group of variables after a cost shock under the optimal Taylor rule. The red dashed line shows the results for the case of the full information rational expectations model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts, and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume that agents use the expectations operator $\mathbb{E}_t^{nh}[\pi_{t+1}] = \mathbb{E}_t^{KF}[\pi_{t+1}] + \theta\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{t-3}^{KF}[\pi_{t+1}]\right)$. Horizontal axis denotes months after the shock.

6.2 Optimal response to a taste shock

Figure 13 shows the impulse response functions to a taste shock and an optimal response from the central bank.[20] As it is well known in the literature, upon a taste shock, the optimal response of the central bank is to strongly raise the interest rate. What follows is that the central bank manages to bring down both the output gap and the inflation rate to their steady state values.

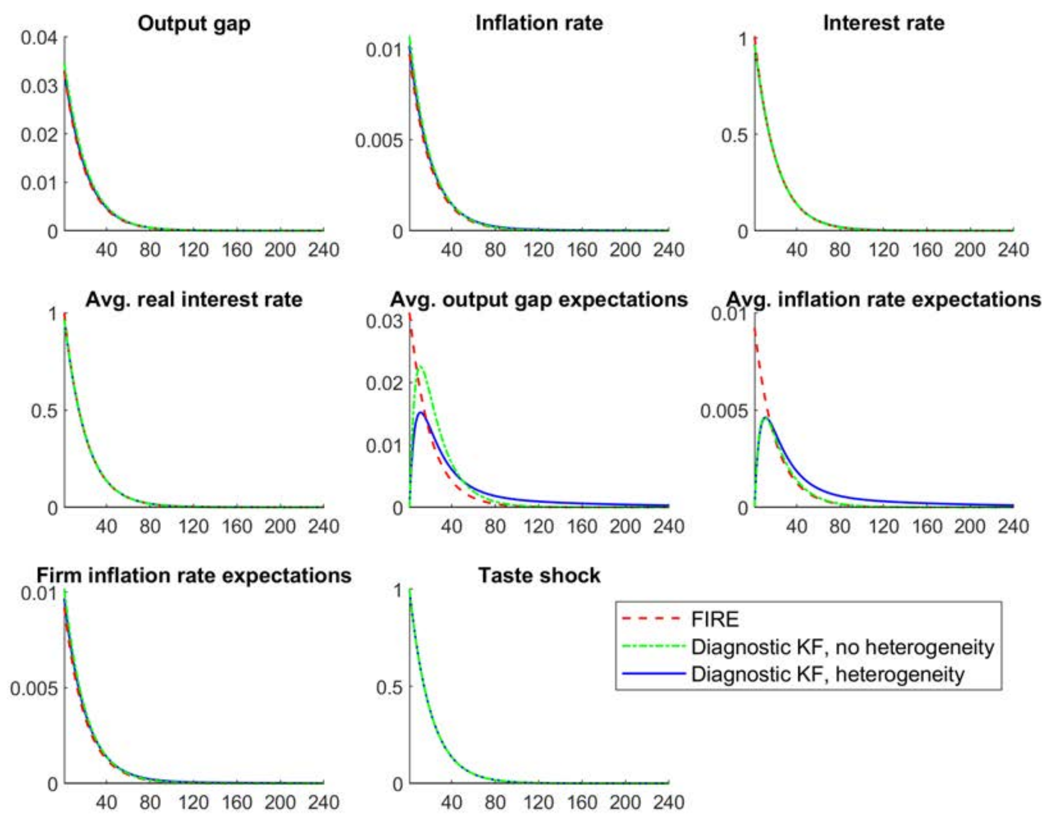
We find no significant difference in the response of the central bank between the FIRE and diagnostic cases. While in the baseline case from Figure 10 the central bank opted to be less active in the diagnostic cases, the optimal Taylor rule case says that the central bank should

[20]Table A.2 in Appendix A shows the optimal parameters.

always be active when facing a taste shock, no matter the type of expectations agents have. In this way, the central bank is able to close the output gap and lower the inflation rate more quickly than with the baseline results.

With the active stance recommended by the optimal Taylor rule, both the output gap and inflation expectations are positive but very close to zero, in both the FIRE and diagnostic cases.

Figure 13: Impulse response functions, optimal Taylor rule, taste shock



Note: Figure shows impulse response functions for a selected group of variables after a taste shock under the optimal Taylor rule. The red dashed line shows the results for the case of the full information rational expectation model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts, and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume the agents use the expectations operator $\mathbb{E}_t^{nh}[\pi_{t+1}] = \mathbb{E}_t^{KF}[\pi_{t+1}] + \theta\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{t-3}^{KF}[\pi_{t+1}]\right)$. Horizontal axis denotes months after the shock.

7 Analyzing an episode of high inflation

In this section we analyze the behavior of the model after the high-inflation episode of 2021.[21] To do so, we feed the model actual monthly data on the output gap, inflation

[21] In this part we go back to the basic calibration of Table 4.

rates, and interest rates up to December 2021.[22] Afterward, we produce forecasts using the different versions of the model (FIRE, diagnostic expectations with homogeneity, diagnostic expectations with heterogeneity).[23]

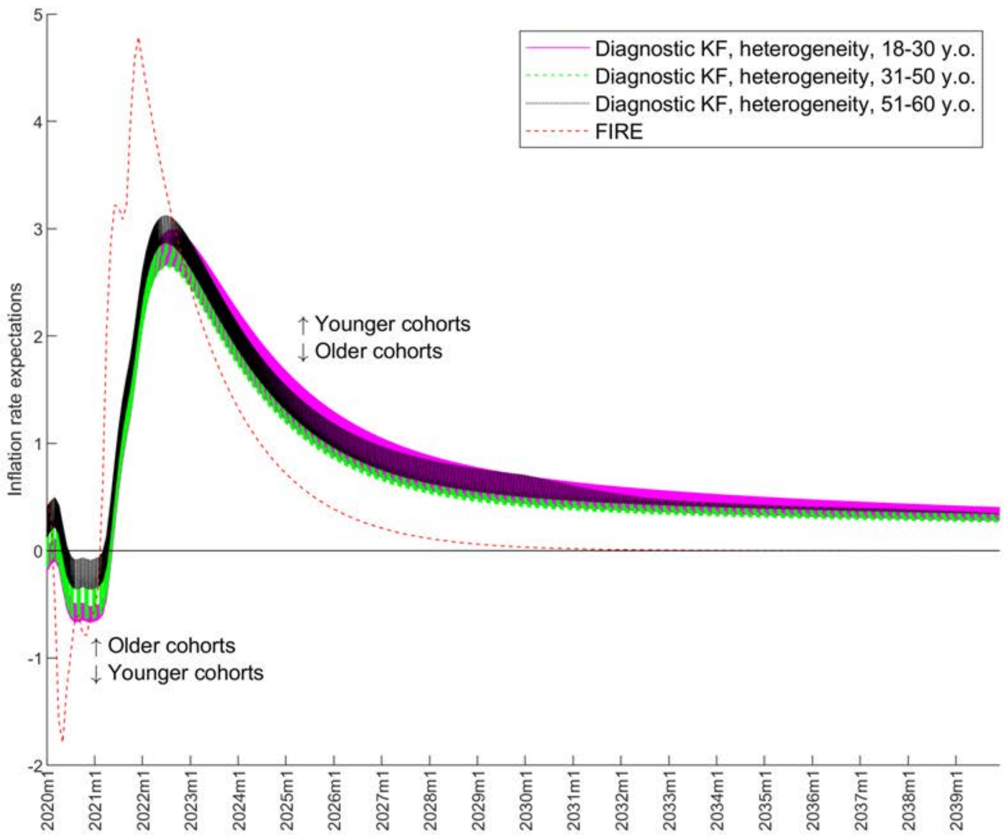
We first analyze the diagnostic expectations with heterogeneity case. Figure 14 shows the inflation rate diagnostic expectations by cohort according to our model and the data. Before 2021 we see that older cohorts had the highest inflation expectations. This is because older cohorts experienced the high-inflation episodes of the 60s, 70s, and early 80s. They are followed by the intermediate cohorts and finally by the youngest cohorts, who experienced low and stable inflation rates from the 90s to the 10s.

However, things change after the high-inflation episode of 2021. Overall, inflation expectations go up after the episode. However, the order of inflation expectations across cohorts changes. The youngest cohorts start having the highest expectations of all, since an important part of their memory is now related to the high-inflation episode. This is related to empirical fact 1 (Figure 1), where we showed that young cohorts started having the highest inflation expectations after 2021.

[22]We use monthly series from March 1967 to December 2021. We go as far as the data allow to build the memory that agents use as a reference. For the output gap we use the National Activity Index (CFNAI) from the Federal Reserve Bank of Chicago. For the interest rate we use the effective federal funds rate. For the inflation rate we use the CPI 12-month percentage change.

[23]We present the shocks that, according to our model, explain the observed data in Figure A.2 in Appendix A.

Figure 14: Impulse response functions, inflation rate diagnostic expectations by cohort, forecast

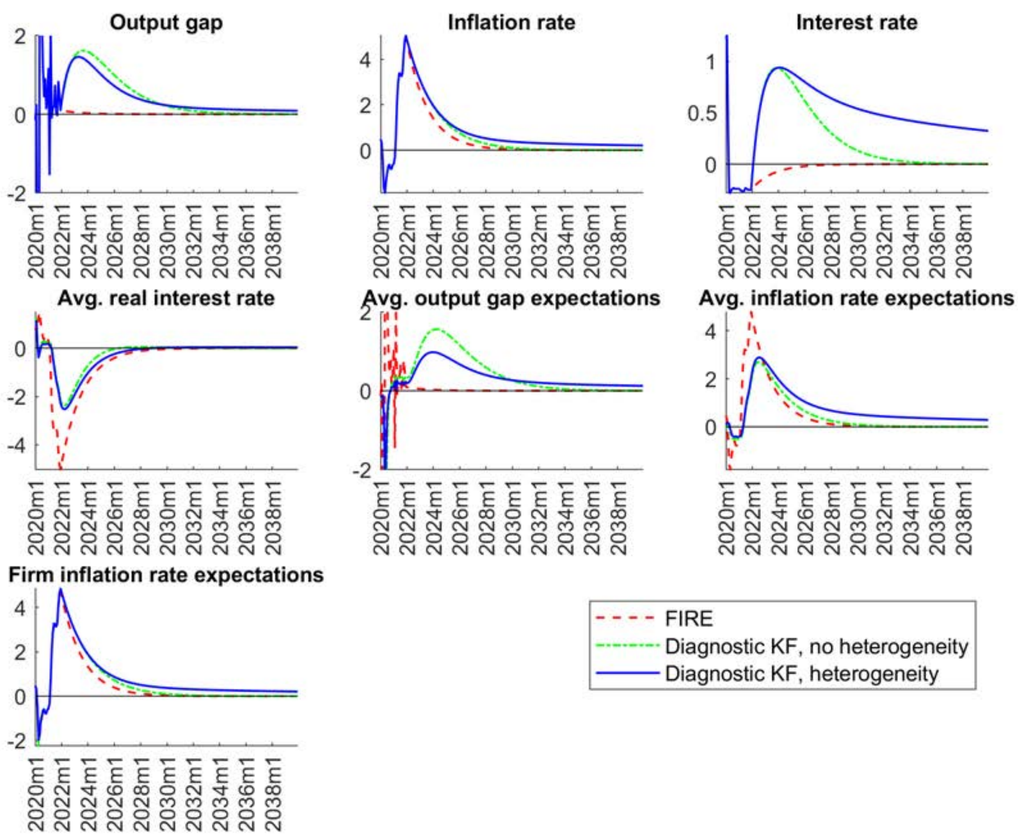


Note: Figure shows the heterogeneous expectations generated by the diagnostic Kalman filter and the data (up to December 2021). Cohorts denote age in 2021. Horizontal axis denotes months.

We now turn to the analysis of all the variables and the comparison across different model specifications. Figure 15 presents how variables evolve based on our model and the data. After 2021, when we allow for diagnostic expectations and heterogeneity, average inflation expectations are higher and more persistent with respect to the other cases. This is because agents remember and anchor their expectations to what they experienced in the past. Then, with diagnostic expectations and heterogeneity the inflation rate is also higher and more persistent than in the other cases. Hence, the central bank must react more strongly.

Another point to notice is that, because of memory, agents remember the high-inflation episode far into the future. In our model inflation expectations remain high even 10 years after the episode. As a consequence, the observed inflation rate and the interest rate also remain high.

Figure 15: Impulse response functions, forecast



Note: Figure shows impulse response functions for a selected group of variables according to the model and the data (up to December 2021). The red dashed line shows the results for the case of the full information rational expectations model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume that agents use the expectations operator $\mathbb{E}_t^{nh}\left[\pi_{t+1}\right] = \mathbb{E}_t^{KF}\left[\pi_{t+1}\right] + \theta\left(\mathbb{E}_t^{KF}\left[\pi_{t+1}\right] - \mathbb{E}_{t-3}^{KF}\left[\pi_{t+1}\right]\right)$. Horizontal axis denotes months.

8 Conclusions

This paper studies the macroeconomic consequences of heterogeneous inflation expectations. We first show that inflation expectations are heterogeneous across cohorts. Based on Bordalo et al. (2020), we introduce a Kalman filter augmented with diagnostic expectations to model the inflation forecast formation process. We structurally estimate the relevant diagnostic parameter, concluding that individuals effectively consider their past inflation histories when forecasting. In their predictions, they under-weight recent developments in the inflation series, adding positive weight to their prior experiences. With this anchoring-to-the-past mechanism in the expectations-formation process, we can effectively replicate a relevant part of the heterogeneity across cohorts and time observed in the inflation expectations data.

Based on Bianchi, Ilut, and Saijo (2021) and L'Huillier, Singh, and Yoo (2021), we then incorporate the idea of heterogeneous inflation expectations into a New Keynesian model that includes diagnostic expectations and heterogeneous cohorts. The model suggests that heterogeneous expectations are essential in the economy's response to supply and demand shocks. Heterogeneous expectations anchor the output gap's aggregate response to agents' memories while increasing the persistence of the effect of the shocks.

We then perform an optimal Taylor rule exercise where the central bank seeks to minimize the variance of the output gap and the inflation rate. Under heterogeneity in expectations, the optimal response of the central bank after a negative supply shock or a positive demand shock is to be more active in controlling inflation, as agents have a long memory and remember current shocks far into the future. In this way, the central bank prevents inflation from rising and, more importantly, prevents agents from incorporating high-inflation episodes into their memories.

These results have relevant implications for the current macroeconomic environment. The model suggests that the 2021 high-inflation episode, even though it may be transitory, could have long-lasting effects: new cohorts will incorporate the high-inflation episode into their histories of inflation, adjusting future expectations. For example, cohorts from the US that experienced high inflation during the 70s hold higher expectations to this day, even after going through the low-inflation environment characteristic of the last several decades. Allowing high inflation today produces higher and more persistent inflation expectations in the future as people carry their beliefs over time.

Our results have important policy implications. One main takeaway from this paper is that heterogeneous inflation expectations anchor the aggregate response of inflation and the output gap to the agents' memory, increasing the persistence of the effects of the shocks. Therefore, the optimal response of monetary authorities when inflation starts rising is to take an active stance. An energetic response prevents current high inflation and prevent agents from incorporating high-inflation episodes into their memories, thus preventing higher future inflation expectations.

References

Adam, Klaus. 2003. "Learning and equilibrium selection in a monetary overlapping generations model with sticky prices." *Review of Economic Studies* 70 (4):887–907. URL https://doi.org/10.1111/1467-937X.00271.

Afrouzi, Hassan. 2020. "Strategic inattention, inflation dynamics, and the non-neutrality of money." URL http://dx.doi.org/10.2139/ssrn.3576296.

Airaudo, Marco, Salvatore Nisticò, and Luis-Felipe Zanna. 2015. "Learning, monetary policy, and asset prices." *Journal of Money, Credit and Banking* 47 (7):1273–1307. URL https://www.jstor.org/stable/24500915.

Bianchi, Francesco, Cosmin L Ilut, and Hikaru Saijo. 2021. "Implications of diagnostic expectations: Theory and applications." URL https://doi.org/10.3386/w28604.

Blanchard, Olivier J. 1985. "Debt, deficits, and finite horizons." *Journal of Political Economy* 93 (2):223–247. URL https://doi.org/10.1086/261297.

Blanchard, Olivier J, Jean-Paul L'Huillier, and Guido Lorenzoni. 2013. "News, noise, and fluctuations: An empirical exploration." *American Economic Review* 103 (7):3045–70. URL https://doi.org/10.1257/aer.103.7.3045.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in macroeconomic expectations." *American Economic Review* 110 (9):2748–82. URL https://doi.org/10.1257/aer.20181219.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic expectations and stock returns." *Journal of Finance* 74 (6):2839–2874. URL https://doi.org/10.1111/jofi.12833.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2018. "Diagnostic expectations and credit cycles." *Journal of Finance* 73 (1):199–227. URL https://doi.org/10.1111/jofi.12586.

Branch, William A. 2004. "The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations." *Economic Journal* 114 (497):592–621. URL https://doi.org/10.1111/j.1468-0297.2004.00233.x.

Branch, William A and George W Evans. 2011. "Monetary policy and heterogeneous expectations." *Economic Theory* 47 (2):365–393. URL https://doi.org/10.1007/s00199-010-0539-9.

Branch, William A and Bruce McGough. 2009. "A New Keynesian model with heterogeneous expectations." *Journal of Economic Dynamics and Control* 33 (5):1036–1051. URL https://doi.org/10.1016/j.jedc.2008.11.007.

———. 2010. "Dynamic predictor selection in a new Keynesian model with heterogeneous expectations." *Journal of Economic Dynamics and Control* 34 (8):1492–1508. URL https://doi.org/10.1016/j.jedc.2010.03.012.

———. 2018. "Heterogeneous expectations and micro-foundations in macroeconomics." *Handbook of computational economics* 4:3–62. URL https://doi.org/10.1016/bs.hescom.2018.03.001.

Brazier, Alex, Richard Harrison, Mervyn King, and Anthony Yates. 2008. "The Danger of Inflating Expectations of Macroeconomic Stability: Heuristic Switching in an Overlapping-Generations Monetary Model." *International Journal of Central Banking* 4 (2):219–254.

URL https://www.ijcb.org/journal/ijcb08q2a6.htm.

Brock, William A and Cars H Hommes. 1997. "A rational route to randomness." *Econometrica* 65 (5):1059–1095. URL https://doi.org/10.2307/2171879.

Campos, Chris, Michael McMain, and Mathieu Pedemonte. 2022. "Understanding Which Prices Affect Inflation Expectations." *Economic Commentary* (2022-06). URL https://www.doi.org/10.26509/frbc-ec-202206.

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko. 2022. "The macroeconomic expectations of firms." URL https://doi.org/10.3386/w30042.

Carroll, Christopher D. 2003. "Macroeconomic expectations of households and professional forecasters." *Quarterly Journal of Economics* 118 (1):269–298. URL https://doi.org/10.1162/00335530360535207.

Carton, Benjamin. 2012. "Monetary-Policy Tradeoff in Overlapping Generations DSGE Models." .

Cheung, Yin-Wong. 1993. "Exchange rate risk premiums." *Journal of International Money and Finance* 12 (2):182–194. URL https://doi.org/10.1016/0261-5606(93)90023-5.

Coibion, Olivier. 2006. "Inflation inertia in sticky information models." *Contributions to Macroeconomics* 6 (1):1–29. URL https://doi.org/10.2202/1534-6005.1374.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Maarten Van Rooij. 2019. "How does consumption respond to news about inflation? Field evidence from a randomized control trial." URL https://doi.org/10.3386/w26106.

Coibion, Olivier and Yuriy Gorodnichenko. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review* 105 (8):2644–78. URL https://www.doi.org/10.1257/aer.20110306.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. "How do firms form their expectations? New survey evidence." *American Economic Review* 108 (9):2671–2713. URL https://doi.org/10.1257/aer.20151299.

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele. 2020. "Inflation expectations and firm decisions: New causal evidence." *Quarterly Journal of Economics* 135 (1):165–219. URL https://doi.org/10.1093/qje/qjz029.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A Lochstoer. 2017. "Asset pricing when 'this time is different'." *Review of Financial Studies* 30 (2):505–535. URL https://doi.org/10.1093/rfs/hhw084.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber. 2021. "Exposure to grocery prices and inflation expectations." *Journal of Political Economy* 129 (5):1615–1639. URL https://doi.org/10.1086/713192.

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber. 2022. "What Do the Data Tell Us About Inflation Expectations?" URL https://doi.org/10.3386/w29825.

De Grauwe, Paul. 2011. "Animal spirits and monetary policy." *Economic Theory* 47 (2):423–457. URL https://www.jstor.org/stable/41485845.

De Grauwe, Paul and Yuemei Ji. 2020. "Structural reforms, animal spirits, and monetary policies." *European Economic Review* 124:103395. URL https://doi.org/10.1016/j.euroecorev.2020.103395.

Di Bartolomeo, Giovanni, Marco Di Pietro, and Bianca Giannini. 2016. "Optimal monetary policy in a New Keynesian model with heterogeneous expectations." *Journal of Economic Dynamics and Control* 73:373–387. URL https://doi.org/10.1016/j.jedc.2016.10.003.

Dietrich, Alexander M. 2022. "Consumption Categories, Household Attention, and Inflation Expectations: Implications for Optimal Monetary Policy." .

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen. 2018. "Asset prices and portfolio choice with learning from experience." *Review of Economic Studies* 85 (3):1752–1780. URL https://doi.org/10.1093/restud/rdx077.

Evans, George W and Seppo Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press. URL https://doi.org/10.1515/9781400824267.

Evans, George W, Seppo Honkapohja, and Ramon Marimon. 2001. "Convergence in monetary inflation models with heterogeneous learning rules." *Macroeconomic Dynamics* 5 (1):1–31. URL https://doi.org/10.1017/S1365100501018016.

Fujiwara, Ippei and Yuki Teranishi. 2008. "A dynamic new Keynesian life-cycle model: Societal aging, demographics, and monetary policy." *Journal of Economic Dynamics and Control* 32 (8):2398–2427. URL https://doi.org/10.1016/j.jedc.2007.09.002.

Gabaix, Xavier. 2020. "A behavioral New Keynesian model." *American Economic Review* 110 (8):2271–2327. URL https://doi.org/10.1257/aer.20162005.

Gali, Jordi. 2015. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*. Princeton University Press.

———. 2021. "Monetary policy and bubbles in a new keynesian model with overlapping generations." *American Economic Journal: Macroeconomics* 13 (2):121–67. URL https://www.aeaweb.org/articles?id=10.1257/mac.20180427.

Gasteiger, Emanuel. 2014. "Heterogeneous expectations, optimal monetary policy, and the merit of policy inertia." *Journal of Money, Credit and Banking* 46 (7):1535–1554. URL https://doi.org/10.1111/jmcb.12149.

Gáti, Laura. 2020. "Monetary Policy & Anchored Expectations: An Endogenous Gain Learning Model." .

Gorodnichenko, Yuriy and Dmitriy Sergeyev. 2021. "Zero lower bound on inflation expectations." URL https://doi.org/10.3386/w29496.

Hajdini, Ina, Edward S Knotek II, John Leer, Mathieu Pedemonte, Robert Rich, and Raphael Schoenle. 2022a. "Indirect Consumer Inflation Expectations: Theory and Evidence." *Federal Reserve Bank of Cleveland Working Paper* URL https://doi.org/10.26509/frbc-wp-202235.

Hajdini, Ina, Edward S. Knotek II, John Leer, Mathieu Pedemonte, Robert W. Rich, and Raphael S. Schoenle. 2022b. "Low Passthrough from Inflation Expectations to Income Growth Expectations: Why People Dislike Inflation." Working Paper 22-21, Federal Reserve Bank of Cleveland. URL https://doi.org/10.26509/frbc-wp-202221.

Hajdini, Ina, Edward S Knotek II, Mathieu Pedemonte, Robert Rich, John Leer, and Raphael Schoenle. 2022c. "Indirect Consumer Inflation Expectations." *Economic Commentary* (2022-03). URL https://doi.org/10.26509/frbc-ec-202203.

Hommes, Cars. 2021. "Behavioral and experimental macroeconomics and policy analysis: A complex systems approach." *Journal of Economic Literature* 59 (1):149–219. URL https://doi.org/10.1257/jel.20191434.

Hürtgen, Patrick. 2014. "Consumer misperceptions, uncertain fundamentals, and the business cycle." *Journal of Economic Dynamics and Control* 40:279–292. URL https://doi.org/10.1016/j.jedc.2014.01.008.

Jump, Robert Calvert, Cars Hommes, and Paul Levine. 2019. "Learning, heterogeneity, and complexity in the New Keynesian model." *Journal of Economic Behavior & Organization* 166:446–470. URL https://doi.org/10.1016/j.jebo.2019.07.014.

Jump, Robert Calvert and Paul Levine. 2019. "Behavioural new Keynesian models." *Journal of Macroeconomics* 59:59–77. URL https://doi.org/10.1016/j.jmacro.2018.11.002.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. "Monetary policy according to HANK." *American Economic Review* 108 (3):697–743. URL https://doi.org/10.1257/aer.20160042.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2020. "The tail that wags the economy: Beliefs and persistent stagnation." *Journal of Political Economy* 128 (8):2839–2879. URL https://doi.org/10.1086/707735.

Kuchler, Theresa and Basit Zafar. 2019. "Personal experiences and expectations about aggregate outcomes." *Journal of Finance* 74 (5):2491–2542. URL https://doi.org/10.1111/jofi.12819.

L'Huillier, Jean-Paul, Sanjay R Singh, and Donghoon Yoo. 2021. "Diagnostic expectations and macroeconomic volatility." .

Madeira, Carlos and Basit Zafar. 2015. "Heterogeneous inflation expectations and learning." *Journal of Money, Credit and Banking* 47 (5):867–896. URL https://doi.org/10.1111/jmcb.12230.

Malmendier, Ulrike. 2021. "Exposure, experience, and expertise: Why personal histories matter in economics." *Journal of the European Economic Association* 19:2857–2894. URL https://doi.org/10.1093/jeea/jvab045.

Malmendier, Ulrike and Stefan Nagel. 2016. "Learning from inflation experiences." *Quarterly Journal of Economics* 131 (1):53–87. URL https://doi.org/10.1093/qje/qjv037.

Malmendier, Ulrike, Stefan Nagel, and Zhen Yan. 2021. "The making of hawks and doves." *Journal of Monetary Economics* 117:19–42. URL https://doi.org/10.1016/j.jmoneco.2020.04.002.

Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco. 2020a. "Investor experiences and financial market dynamics." *Journal of Financial Economics* 136 (3):597–622. URL https://doi.org/10.1016/j.jfineco.2019.11.002.

———. 2020b. "Investor experiences and international capital flows." *Journal of International Economics* 124:103302. URL https://doi.org/10.1016/j.jinteco.2020.103302.

Malmendier, Ulrike and Jessica A Wachter. 2022. "Memory of Past Experiences and Economic Decisions." .

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. URL https://doi.org/10.1162/003355302320935034.

———. 2006. "Pervasive stickiness." *American Economic Review* 96 (2):164–169. URL https://doi.org/10.1257/000282806777211937.

———. 2007. "Sticky information in general equilibrium." *Journal of the European Economic Association* 5 (2-3):603–613. URL https://doi.org/10.1162/jeea.2007.5.2-3.603.

———. 2011. "Imperfect information and aggregate supply." *Handbook of Monetary Economics* 3:183–229. URL https://doi.org/10.1016/B978-0-444-53238-1.00005-3.

Marcet, Albert and Thomas J Sargent. 1989. "Convergence of least squares learning mechanisms in self-referential linear stochastic models." *Journal of Economic Theory* 48 (2):337–368. URL https://doi.org/10.1016/0022-0531(89)90032-X.

Massaro, Domenico. 2013. "Heterogeneous expectations in monetary DSGE models." *Journal of Economic Dynamics and Control* 37 (3):680–692. URL https://doi.org/10.1016/j.jedc.2012.11.001.

Nagel, Stefan and Zhengyang Xu. 2022. "Asset pricing with fading memory." *Review of Financial Studies* 35–5:2190–2245. URL https://doi.org/10.1093/rfs/hhab086.

Nimark, Kristoffer P. 2014. "Man-bites-dog business cycles." *American Economic Review* 104 (8):2320–67. URL https://doi.org/10.1257/aer.104.8.2320.

Orphanides, Athanasios and John C Williams. 2007. "Robust monetary policy with imperfect knowledge." *Journal of Monetary Economics* 54 (5):1406–1435. URL https://www.jstor.org/stable/40005102.

Pfajfar, Damjan. 2013. "Formation of rationally heterogeneous expectations." *Journal of Economic Dynamics and Control* 37 (8):1434–1452. URL https://doi.org/10.1016/j.jedc.2013.03.012.

Pivetta, Frederic and Ricardo Reis. 2007. "The persistence of inflation in the United States." *Journal of Economic Dynamics and Control* 31 (4):1326–1358. URL https://doi.org/10.1016/j.jedc.2006.05.001.

Reis, Ricardo. 2009. "Optimal monetary policy rules in an estimated sticky-information model." *American Economic Journal: Macroeconomics* 1 (2):1–28. URL https://doi.org/10.1257/mac.1.2.1.

———. 2020. "The People versus the Markets: A Parsimonious Model of Inflation Expectations." .

Roth, Christopher and Johannes Wohlfart. 2020. "How do expectations about the macroeconomy affect personal expectations and behavior?" *Review of Economics and Statistics* 102 (4):731–748. URL https://doi.org/10.1162/rest_a_00867.

Schraeder, Stefanie. 2016. "Information processing and non-Bayesian learning in financial markets." *Review of Finance* 20 (2):823–853. URL https://doi.org/10.1093/rof/rfv002.

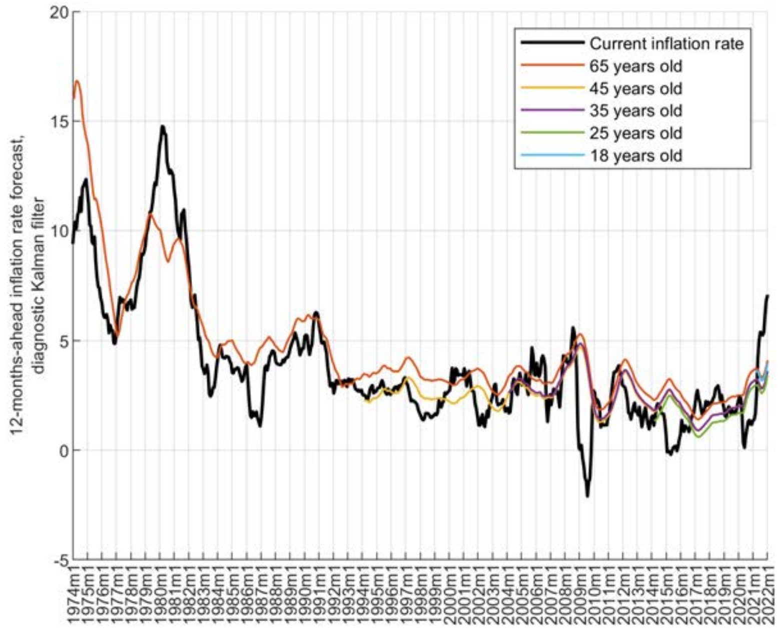
Souleles, Nicholas S. 2004. "Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the Michigan consumer sentiment surveys." *Journal of Money, Credit and Banking* :39–72URL https://www.jstor.org/stable/3839047.

Woodford, Michael. 2001. "Imperfect common knowledge and the effects of monetary policy." URL https://doi.org/10.3386/w8673.

Yaari, Menahem E. 1965. "Uncertain lifetime, life insurance, and the theory of the consumer." *Review of Economic Studies* 32 (2):137–150. URL https://doi.org/10.2307/2296058.

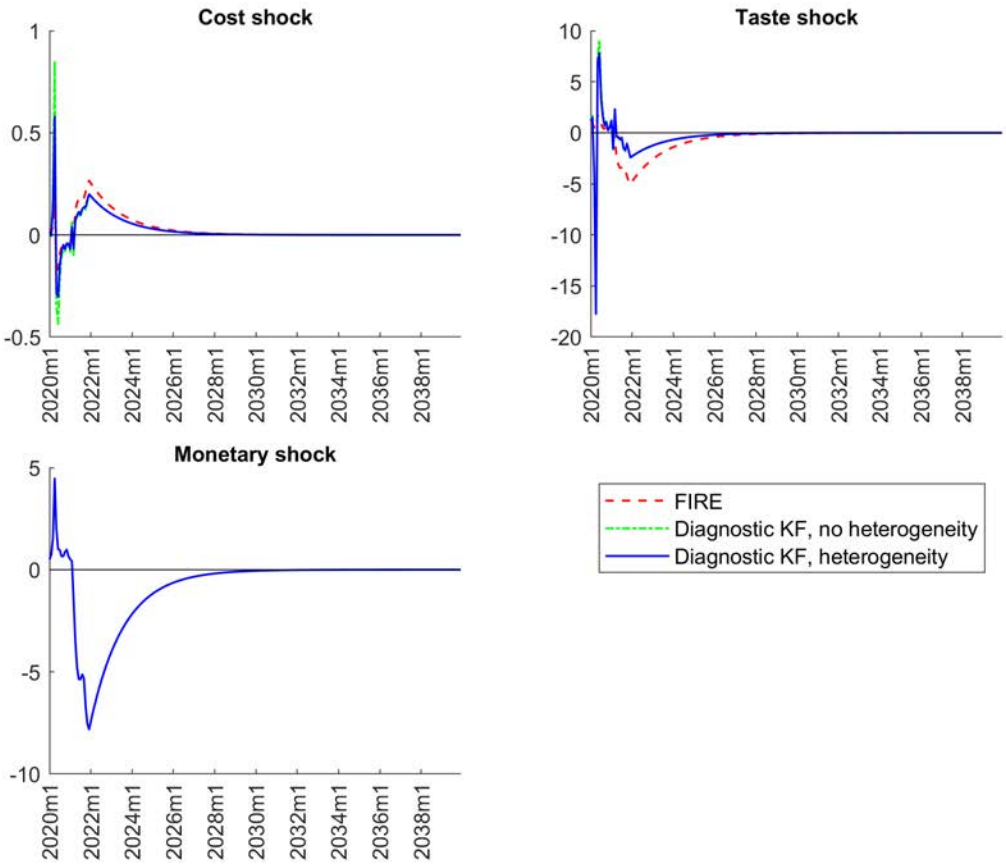
A Additional figures and tables

Figure A.1: Diagnostic Kalman-filter-based inflation forecasts by cohort, full sample



Note: Figure shows forecasts for selected cohorts according to the Kalman-filter-augmented expectations and considering the estimate for θ from Column 1 of Table 2. Selected cohorts differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

Figure A.2: Impulse response functions, shocks, forecast



Note: Figure shows the paths shocks follow according to the model and the data (up to December 2021). The red dashed line shows the results for the case of the full information rational expectations model (FIRE), the green dash-dot line shows the results of a diagnostic Kalman filter model with no heterogeneity across cohorts, and the solid blue line shows the diagnostic Kalman filter model with heterogeneity. For the no heterogeneity case we assume agents use the expectations operator $\mathbb{E}_t^{nh}[\pi_{t+1}] = \mathbb{E}_t^{KF}[\pi_{t+1}] + \theta\left(\mathbb{E}_t^{KF}[\pi_{t+1}] - \mathbb{E}_{t-3}^{KF}[\pi_{t+1}]\right)$. Horizontal axis denotes months.

Table A.1: Optimal Taylor rule parameters, cost shock

	χ_π^*	χ_y^*
FIRE	582.33	3.38
Diagnostic KF, no heterogeneity	32.25	0.00
Diagnostic KF, heterogeneity	43.72	0.00

Note: Table shows the Taylor rule parameters that minimize objective function $\mathbb{E}_t\left[\pi_t^2 + \vartheta y_t^2\right]$ when an unexpected cost shock hits the economy.

Table A.2: Optimal Taylor rule parameters, taste shock

	χ_π^*	χ_y^*
FIRE	8.73	28.10
Diagnostic KF, no heterogeneity	8.34	25.37
Diagnostic KF, heterogeneity	9.23	26.46

Note: Table shows the Taylor rule parameters that minimize objective function $\mathbb{E}_t\left[\pi_t^2 + \vartheta y_t^2\right]$ when an unexpected taste shock hits the economy.

B Normality of diagnostic expectations

Let $f\left(\pi_{t+1}|\mathcal{I}_t\right)$ be the true distribution of the future inflation rate conditional on information set $\mathcal{I}_t$. We assume this behaves as

$$f\left(\pi_{t+1}|\mathcal{I}_t\right) \sim N\left(\mathbb{E}_t^{KF}\left[\pi_{t+1}\right], \sigma_\pi^2\right),$$

where $\mathbb{E}_t^{KF}\left[\pi_{t+1}\right]$ is the expectation according to the standard Kalman filter and σ_π^2 is the variance. We further assume this distribution to be true across all cohorts i.

Let $f\left(\pi_{t+1}|\mathcal{I}_{i,t}^{ref}\right)$ be the distribution of the inflation rate conditional on the referential information set for cohort i. This distribution behaves as

$$f\left(\pi_{t+1}|\mathcal{I}_{i,t}^{ref}\right) \sim N\left(\mathbb{E}_{i,t}^{ref}\left[\pi_{t+1}\right], \sigma_\pi^2\right),$$

where

$$\mathbb{E}_{i,t}^{ref}\left[\pi_{t+1}\right] = \frac{\sum_{j=1}^{t-k_i} \mathbb{E}_{i,t-j}^{KF}\left[\pi_{t+1}\right]}{t-k_i}.$$

Given these two elements, we define the diagnostic distribution as

$$f_{i,t}^{\theta}\left(\pi_{t+1}\right) = f\left(\pi_{t+1}|\mathcal{I}_t\right) D_{i,t}^{\theta}\left(\pi_{t+1}\right),$$

with

$$D_{i,t}^{\theta}\left(\pi_{t+1}\right) = \left[\frac{f\left(\pi_{t+1}|\mathcal{I}_t\right)}{f\left(\pi_{t+1}|\mathcal{I}_{i,t}^{ref}\right)^{t-k_i}}\right]^{\theta} Z_{i,t},$$

where $Z_{i,t}$ is a term that ensures that $f_{i,t}^{\theta}\left(\pi_{t+1}\right)$ integrates to 1.

Therefore, the pdf of the diagnostic distribution is[24]

$$f_{i,t}^{\theta}\left(\pi_{t+1}\right) = \frac{\left[\frac{1}{\sigma_\pi\sqrt{2\pi}} \exp\left\{-\frac{\left(\pi_{t+1}-\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right]\right)^2}{2\sigma_\pi^2}\right\}\right]^{(1+\theta)}}{\left[\frac{1}{\sigma_\pi\sqrt{2\pi}} \exp\left\{-\frac{\left(\pi_{t+1}-\mathbb{E}_{i,t}^{ref}\left[\pi_{t+1}\right]\right)^2}{2\sigma_\pi^2}\right\}\right]^{\theta}} Z_{i,t},$$

[24]Note that π_{t+1} denotes the future inflation rate, while π denotes the constant equal to 3.14.

where we define $Z_{i,t}^{-1} = \int f_{i,t}^{\theta}\left(\pi_{t+1}\right) d\pi_{t+1}$.

We can make the following approximation:

$$f_{i,t}^{\theta}\left(\pi_{t+1}\right) \approx \frac{1}{\sigma_{\pi}\sqrt{2\pi}} \exp\left\{-\frac{\left(\pi_{t+1} - \mathbb{E}_{i,t}^{\theta}\left[\pi_{t+1}\right]\right)^{2}}{2\sigma_{\pi}^{2}}\right\} Z_{i,t},$$

where

$$\mathbb{E}_{i,t}^{\theta}\left[\pi_{t+1}\right] = \mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] + \theta\left(\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] - \mathbb{E}_{i,t}^{ref}\left[\pi_{t+1}\right]\right).$$

Thus, we conclude that

$$f_{i,t}^{\theta}\left(\pi_{t+1}\right) \sim N\left(\mathbb{E}_{i,t}^{\theta}\left[\pi_{t+1}\right], \sigma_{\pi}^{2}\right).$$

C Diagnostic Kalman filter with AR(1) assumption

In this section we repeat the forecasting exercise from Section 4 but replacing the random walk assumption with an AR(1) specification. Therefore, agents assume that inflation behaves as

$$\pi_{t+1} = \rho_\pi \pi_t + \varepsilon_t,$$

where the coefficient $\rho_\pi \in [0, 1]$ captures the mean-reversion of the inflation variable. Here, we assume the inflation rate has been properly demeaned.

As before, we assume that the signal is given by

$$s_t = \zeta \pi_{t+1} + \upsilon_t.$$

The forecasted value of the inflation variable is

$$\mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right] = (1 - \zeta K)\, \mathbb{E}_{i,t-1}^{KF}\left[\pi_{t+1}\right] + K s_t,$$

where the difference now lies in the fact that agents use the AR(1) assumption to forecast the inflation rate such that

$$\mathbb{E}_{i,t}^{KF}\left[\pi_{t+h}\right] = \rho_\pi^{h-1} \mathbb{E}_{i,t}^{KF}\left[\pi_{t+1}\right].$$

In this section we assume $\zeta = 1$, $\rho = 0.99$, $\sigma_\varepsilon = 0.15$, $\sigma_\nu = 4.09$ and $\sigma_{\varepsilon\nu} = -0.06$. This gives $K = 0.1744$.

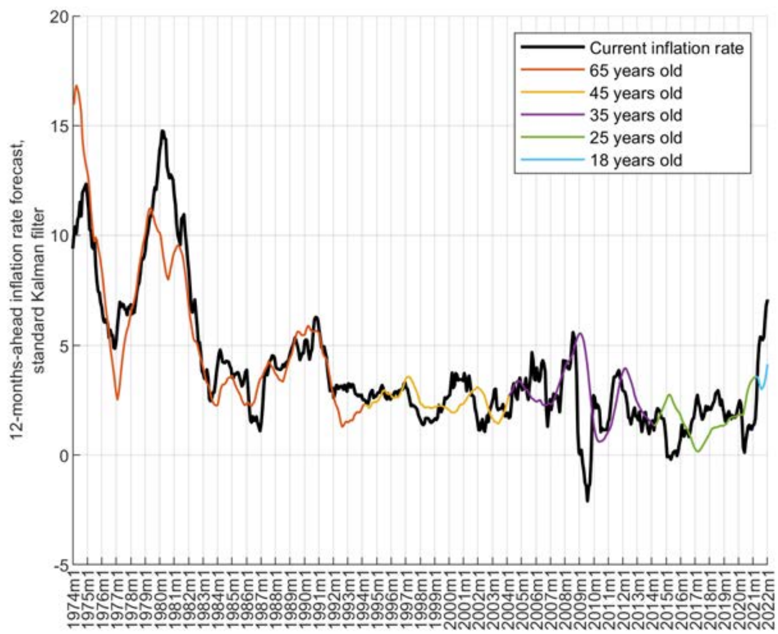
For the different cohorts Figure A.3 presents the standard Kalman filter forecast, while Figure A.4 presents the reference.[25]

Table A.3 presents the result of the diagnostic parameter estimation. In this case, $\theta = -0.809$. Armed with this coefficient, Figure A.5 shows the heterogeneous diagnostic forecasts across cohorts.

Finally, Figure A.6 presents the comparison of the diagnostic forecasts with the AR(1) assumption and the observed forecasts in the data. We find that, as with the random walk process, this version of the diagnostic forecast based on an AR(1) assumption provides a good fit to the data.

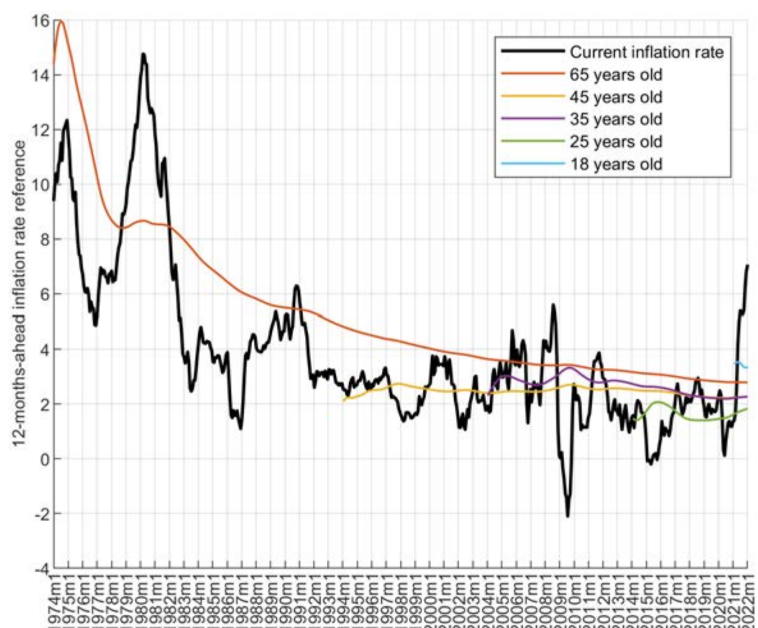
[25]We return the mean to the data before plotting the graphs, where the long-run mean of the inflation rate is 2 percent.

Figure A.3: Standard Kalman-filter-based inflation forecasts by cohort, AR(1)



Note: Figure shows the Kalman filter forecast for the common component for selected cohorts, differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

Figure A.4: Inflation rate reference by cohort, AR(1)



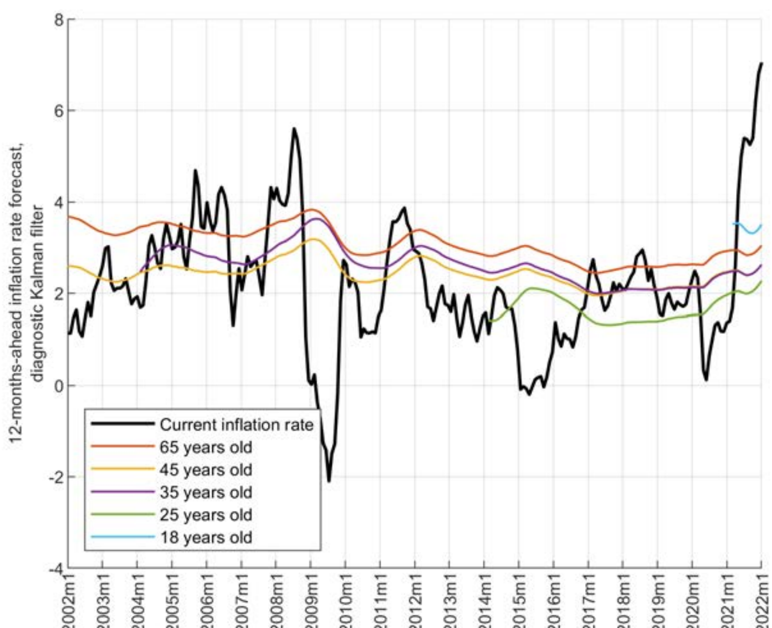
Note: Figure shows the references for selected cohorts obtained according to the Kalman filter and given the history of inflation experienced by the corresponding age group. Selected cohorts are differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

Table A.3: Diagnostic parameter estimation, AR(1)

	(1)
$\mathbb{E}_{i,t}^{ref}\left[\pi_{t+12}\right]$	0.809***
	(0.072)
Time FE	Yes
Observations	101,256
R-squared	0.092

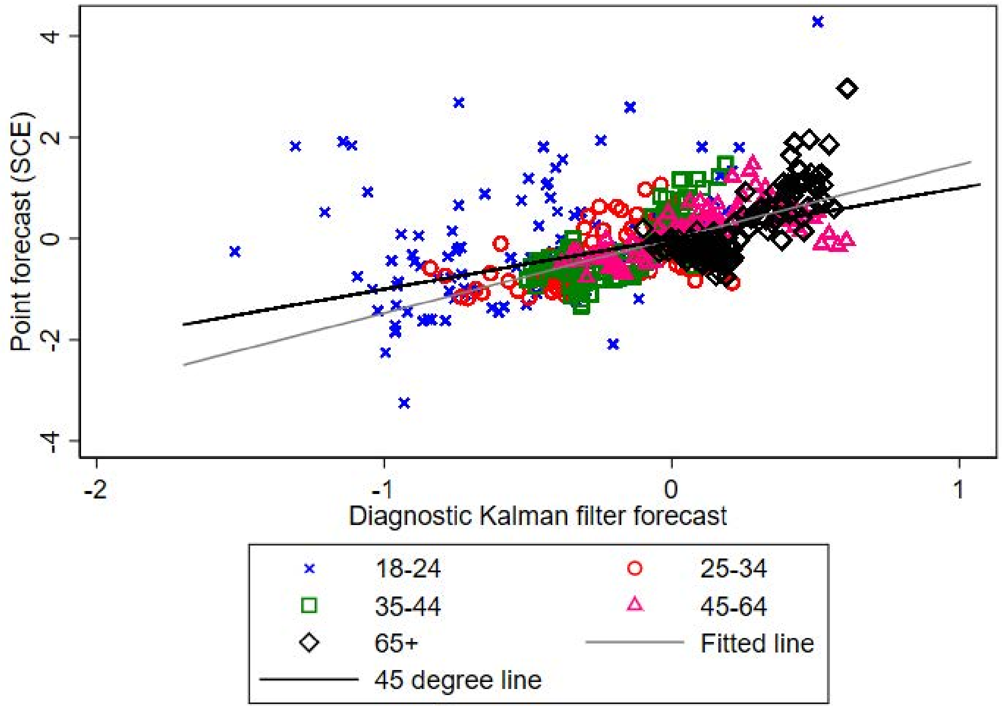
Note: Table shows results of Regression (6). $\mathbb{E}_{i,t}^{ref}\left[\pi_{t+12}\right]$ is the reference constructed for a respondent of age i as explained in the main text. Column (1) has only a time fixed effect as an additional control. Robust standard errors in parentheses. Standard errors clustered by age. Dependent variable trimmed at 10 percent and 90 percent in each period. *** p<0.01, ** p<0.05, * p<0.1.

Figure A.5: Diagnostic Kalman-filter-based inflation forecasts by cohort, AR(1)



Note: Figure shows forecasts for selected cohorts according to the Kalman-filter-augmented expectations and considering the estimate for θ from Column 1 of Table A.3. Selected cohorts differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

Figure A.6: Observed inflation forecasts and diagnostic Kalman filter forecasts, AR(1)



Note: Figure shows binned scatterplot across diagnostic Kalman filter forecasts (x-axis) and point forecast inflation expectations according to the Survey of Consumer Expectations (SCE) of the Federal Reserve Bank of New York (y-axis). Variables demeaned by the intercept. Data go from June 2013 to December 2021. SCE variable trimmed at 10 percent and 90 percent in each period.

D External validity: European data

We check the external validity of our results using data from the Consumer Expectations Survey (CES) of the European Central Bank. It contains monthly data between April 2020 and September 2022 for six countries: Belgium, France, Germany, Italy, the Netherlands, and Spain.[26]

We see that in Europe, lifetime experiences with the inflation rate are also heterogeneous across cohorts, as we show in Figure A.8. Moreover, we see that by 2020 the youngest cohorts had not been exposed to high inflation rates, but this changes after the high inflation rate episode of 2021 and 2022. After this, the youngest cohorts are the ones that show the highest lifetime average for the inflation rate, even larger than that of the people who experienced the high inflation rates of the 80s.

In Figure A.9 we relate the two previous facts and find that in Europe, similar to the US, the larger the inflation rate individuals have experienced in their lifetimes, the higher their inflation expectations.

Table A.4 shows that in Europe, as happened in the US, after controlling for the average lifetime inflation rate, younger generations do not react more strongly to inflation news than older cohorts.

We now turn to the diagnostic Kalman filter of Section 4.1. Using European data we estimate the diagnostic parameter according to Equation 6.[27] In Table A.5 in our baseline specification of Column 1 we find $\theta^{eur} = -0.156$, a parameter that suggests under-reaction to current news. With this parameter, in Figure A.10 we plot inflation expectations according to our diagnostic Kalman filter, across cohorts and in each of the six countries in our sample. We find that the oldest cohorts have the highest inflation expectations before 2021. Then, after 2021 the youngest cohorts start catching up with the oldest ones and even surpass them in some countries.

Table A.5 also shows additional specifications of the estimation of Equation 6. We find that after controlling for cohort and country fixed effects, we still find that agents under-react to current news when forming their expectations. These additional specifications also tell us that the heterogeneity in expectations across cohorts is not due to people of different ages or

[26]There is a relevant difference between the data sets of US and Europe. In the former we have the exact age of the respondents. In the latter we do not have detailed information on the age of the respondents, as they are classified in 4 age groups: 18-34, 35-49, 50-70 and 71+.

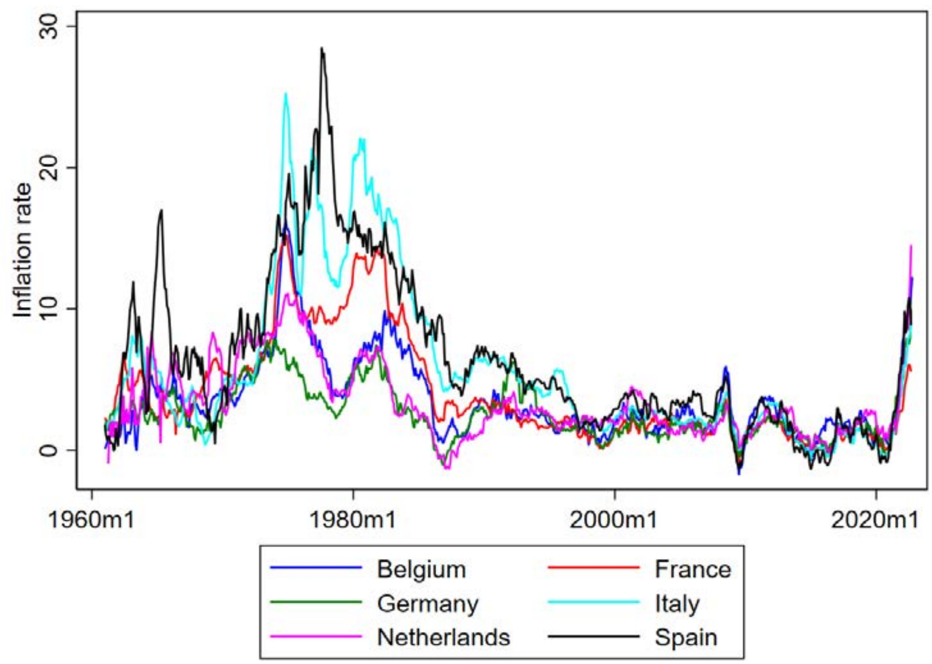
[27]Because we do not know the exact age of the respondents, we do not know which are the exact lifetime average inflation rates they have experienced. Therefore, for this estimation, we assume that every agent in cohort 18-34 has the lifetime average inflation rate of a 25-year-old, every agent in cohort 35-49 has the lifetime average inflation rate of a 35-year-old, every agent in cohort 50-70 has the lifetime average inflation rate of a 50-year-old and every agent in cohort 71+ has the lifetime average inflation rate of a 71-year-old. On the signals used, because the series on the inflation rate of the food component of the CPI have varying starting dates in the different countries, we replace the missing values with the observed inflation rate in order to make the starting dates of all countries uniform.

from different countries facing different consumption bundles or having different preferences, but to the proposed anchoring-to-the-past mechanism. Thus, it is past experiences that define expectations, not the age or the geographic location per se.[28]

Lastly, in Figure A.11 we compare the inflation expectations generated by our diagnostic Kalman filter to the survey data. We see that we have a decent fit to the data.

We conclude that our findings from the main text are also valid for Europe. We find evidence that supports the claim that (i) inflation expectations are also heterogeneous in Europe and (ii) can also be modeled by a diagnostic Kalman filter with under-reaction to current news and over-weighting to the reference term.

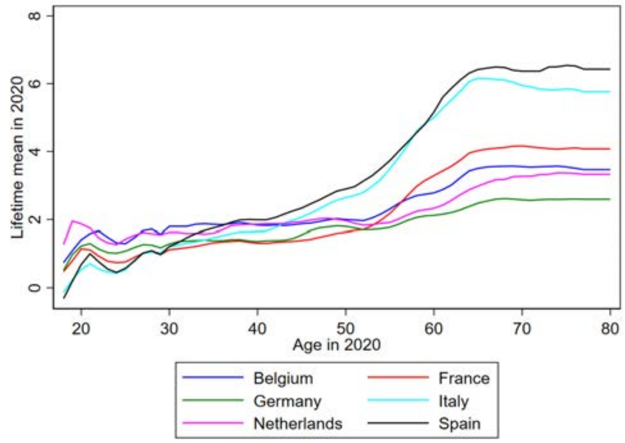
Figure A.7: Inflation rate, Europe



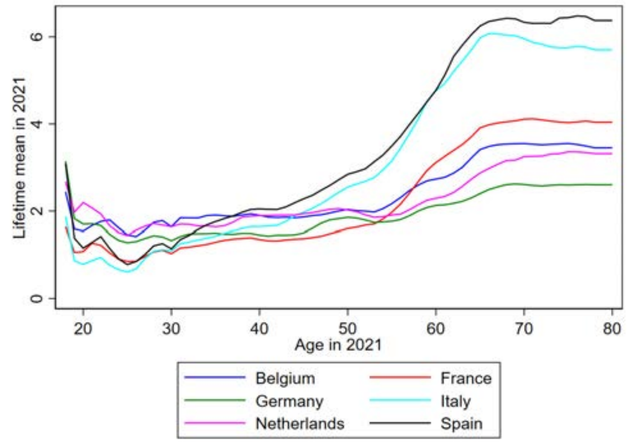
Source: FRED.

[28]See Hajdini et al. (2022a) for a further discussion on this.

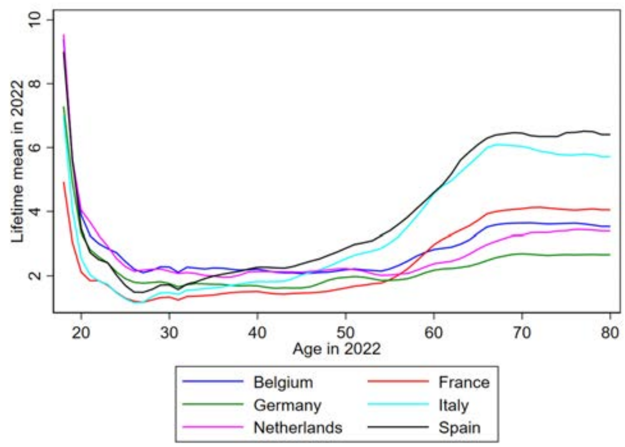
Figure A.8: Lifetime average inflation rate among respondents, Europe



(a) 2020



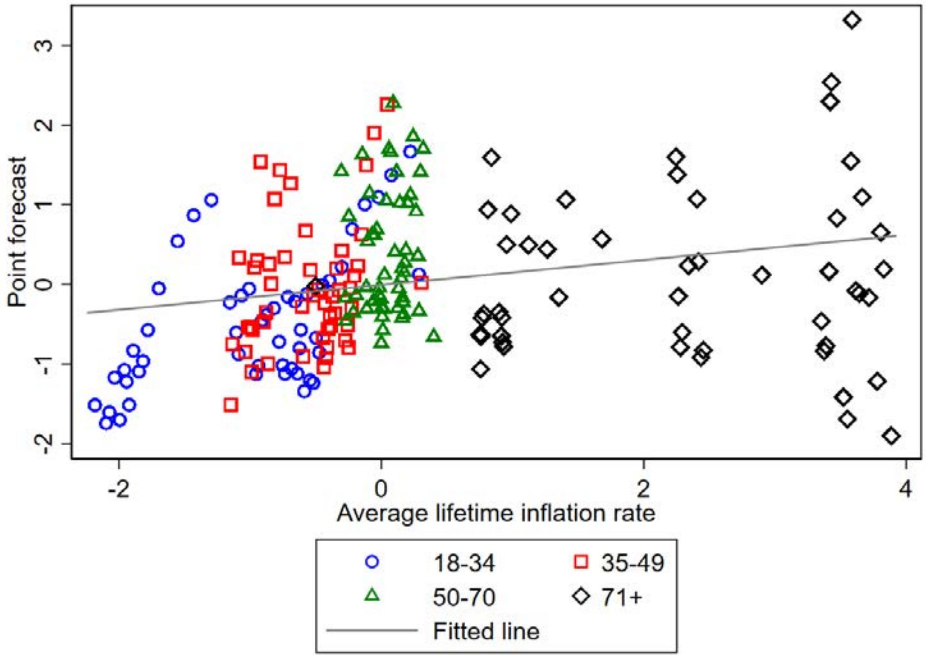
(b) 2021



(c) 2022

Note: Figure shows the mean of the monthly YoY inflation rate that people of the age shown in 2020, 2021, and 2022 have experienced in their lifetimes, starting when they were age 18.
Source: FRED.

Figure A.9: Inflation point forecast and average lifetime inflation rate, Europe



Note: Figure shows binned scatterplot across lifetime average inflation rate bins. Variables residualized by respondent gender and commuting zone. Data go from April 2020 to September 2022. Ages correspond to the interviewee's age at the time of the survey.
Source: Consumer Expectations Survey.

Table A.4: Effects of current and experienced inflation rates on inflation expectations

Dep. var.: Inflation expectations	(1)	(2)	(3)	(4)
Average lifetime inflation rate	0.276*	0.244**	0.301*	0.252*
	(0.132)	(0.100)	(0.149)	(0.129)
Current inflation	0.351***	0.275***		
	(0.033)	(0.057)		
Cohort 35-49		0.155		
		(0.090)		
Cohort 50-70		0.203*		
		(0.112)		
Cohort 71+		-0.401		
		(0.348)		
Current inflation × 35-49		0.074		
		(0.074)		
Current inflation × 50-70		0.122		
		(0.069)		
Current inflation × 71+		0.124		
		(0.083)		
Time FE	No	No	Yes	Yes
Controls	No	No	No	Yes
Observations	294,232	294,232	294,232	294,232
R-squared	0.128	0.135	0.152	0.164

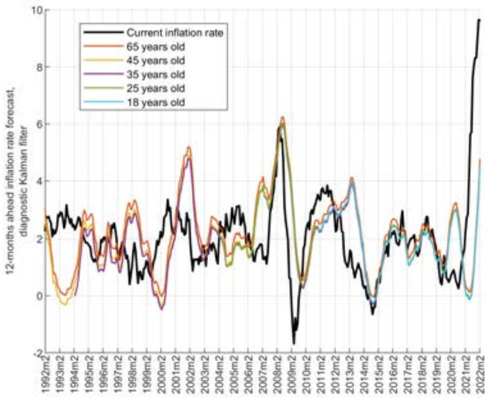
Note: Table shows regressions where the dependent variable is inflation expectations according to the Consumer Expectations Survey (CES) of the European Central Bank. Column (1) shows controls for the average lifetime inflation of respondents of a given age at each period in time and the last inflation measure. Column (2) follows (1) but adds cohort fixed effects and the interaction of those cohort fixed effects with the current inflation. Column (3) follows (1) but adds time fixed effects and, hence, omits the current inflation variable. Column (4) follows (1) but adds time fixed effects and demographic controls. The demographic controls are income, gender, educational level, and country. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered by age. The dependent variable is trimmed, dropping the lower and upper 10 percent of answers in each period.

Table A.5: Diagnostic parameter estimation, Europe

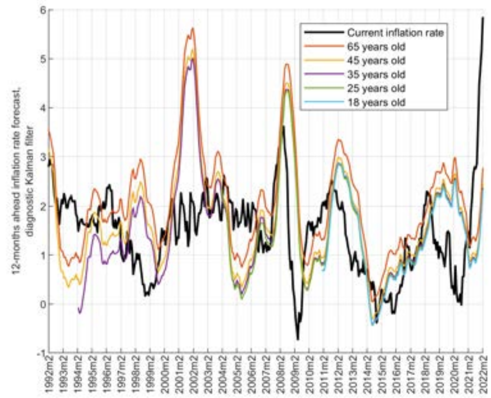
	(1)	(2)	(3)	(4)
$\mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right]$	0.156***	0.208***	0.094***	0.060*
	(0.025)	(0.044)	(0.019)	(0.030)
Time FE	Yes	Yes	Yes	Yes
Controls	No	Cohort	Country	Cohort, country
Observations	271,311	271,311	271,311	271,311
R-squared	0.122	0.130	0.132	0.140

Note: Table shows results of Regression (6), but changing the dependent variable for inflation expectations according to the Consumer Expectations Survey (CES) of the European Central Bank. The independent variable $\mathbb{E}^{ref}_{i,t}\left[\pi_{t+12}\right]$ is the reference constructed for a respondent of age i as explained in the main text. Column (1) considers a time fixed effect as a control. Column (2) has time and cohort fixed effects. Column (3) has time and country fixed effects. Column (4) has time, cohort, and country fixed effects. Standard errors clustered by age in parentheses. Dependent variable trimmed at 10 percent and 90 percent in each period. We use population weights. *** p$<$0.01, ** p$<$0.05, * p$<$0.1.

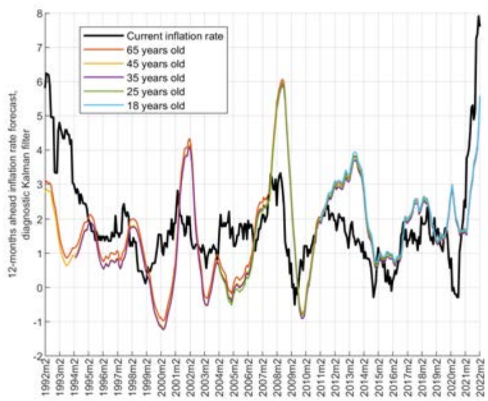
Figure A.10: Diagnostic Kalman-filter-based inflation forecasts by cohort, Europe



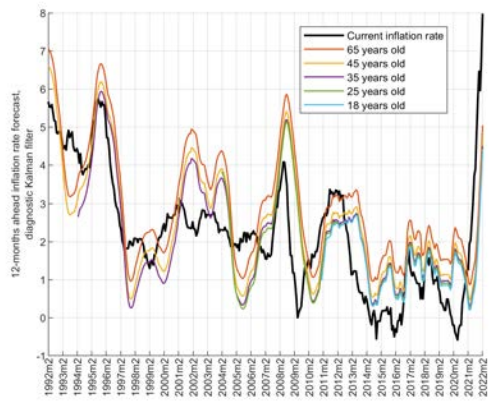
(a) Belgium



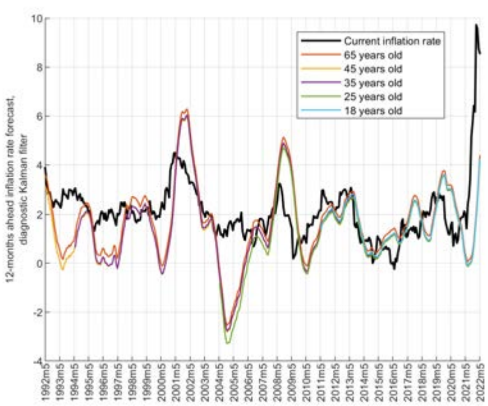
(b) France



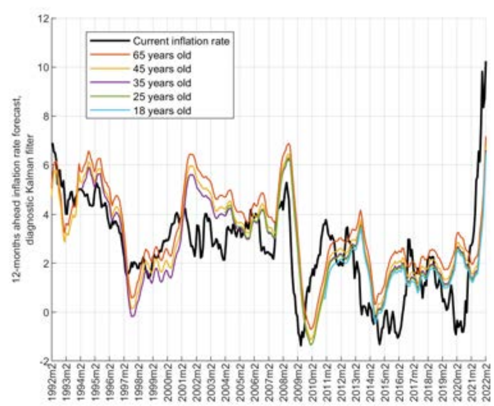
(c) Germany



(d) Italy



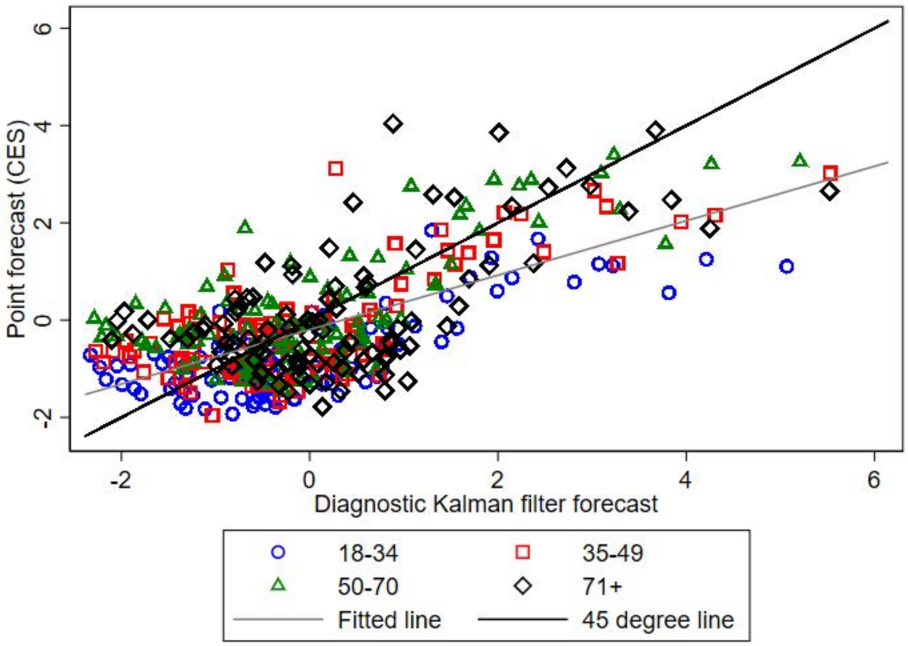
(e) Netherlands



(f) Spain

Note: Figure shows forecasts for selected cohorts according to the Kalman-filter-augmented expectations and considering the estimate for θ from Column 1 of Table A.5. Selected cohorts differentiated by their age in 2021. We further assume that each cohort starts forecasting when they become 18 years old.

Figure A.11: Observed inflation forecasts and diagnostic Kalman filter forecasts, Europe



Note: Figure shows binned scatterplot across diagnostic Kalman filter forecasts (x-axis) and point forecasts of inflation expectations according to the Consumer Expectations Survey (CES) of the European Central Bank (y-axis). Variables demeaned by the intercept. Data go from April 2020 to September 2022. SCE variable trimmed at 10 percent and 90 percent in each period.

E Derivations for firm block

E.1 Final good producer

The final good producer operates in a perfectly competitive market. It produces the final good Y from a CES basket composed by a continuum of intermediate goods $Y(j)$ with $j \in [0,1]$. The maximization problem of this firm is

$$\max_{Y_t(j)} P_t Y_t - \int_0^1 P_t(j)\, Y_t(j)\, dj,$$

subject to

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\varepsilon-1}{\varepsilon}}\, dj \right)^{\frac{\varepsilon}{\varepsilon-1}},$$

where $P_t(j)$ is the price of intermediate good j and ε is the elasticity of substitution in the CES basket.

The first-order condition gives

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\varepsilon} Y_t,$$

which represents the demand for good j.

For the CES basket, we also get a corresponding aggregate price level expression of

$$P_t = \left(\int_0^1 P_t(j)^{1-\varepsilon}\, dj \right)^{\frac{1}{1-\varepsilon}}.$$

E.2 Intermediate good producers

Any given intermediate good producer j will produce the intermediate good $Y(j)$ according to

$$Y_t(j) = A_t L_t(j),$$

where A_t is a process that represents technology and $L_t(j)$ is the labor supplied to firm j. The interemadiate good producer will pay a nominal wage w to workers.

The problem of an intermediate good producer indexed by j is

$$\min w_t L_t(j),$$

subject to

$$Y_t(j) = A_t L_t(j),$$

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\varepsilon} Y_t.$$

The first-order conditions are

$$mc_t = \frac{w_t}{A_t}, \tag{20}$$

where mc is the real marginal cost the firm faces.

E.3 Price setting

We assume that, additionally, intermediate good producers face price rigidities à la Calvo. In any given period, a firm has a probability $1-\phi$ of adjusting its price. That is to say, with probability ϕ this firm will have to keep the price it chose in the previous period. The standard derivation for an optimal reset price results in

$$P_t^* = \frac{\varepsilon}{\varepsilon-1}\frac{X_{1,t}}{X_{2,t}},$$

$$X_{1,t} = \Lambda_t mc_t P_t^{\varepsilon} Y_t + \phi\beta\mathbb{E}_t\left[X_{1,t+1}\right],$$

$$X_{2,t} = \Lambda_t mc_t P_t^{\varepsilon' 1} Y_t + \phi\beta\mathbb{E}_t\left[X_{2,t+1}\right],$$

where P^* is the optimal reset price, X_1 and X_2 are auxiliary variables and $\Lambda{=}u_C(C)$.

The definition of the two auxiliary variables can be rewritten in real terms as

$$x_{1,t} = \Lambda_t mc_t Y_t + \phi\beta\mathbb{E}_t \left[(1+\pi_{t+1})^{\varepsilon} x_{1,t+1}\right], \tag{21}$$

$$x_{2,t} = \Lambda_t Y_t + \phi\beta\mathbb{E}_t \left[(1+\pi_{t+1})^{\varepsilon-1} x_{2,t+1}\right], \tag{22}$$

where

$$x_{1,t} = \frac{X_{1,t}}{P_t^{\varepsilon}},$$

$$x_{2,t} = \frac{X_{2,t}}{P_t^{\varepsilon-1}}.$$

Then, from the reset price definition, we define the reset price inflation rate as

$$(1+\pi_t^*) = \frac{\varepsilon}{\varepsilon-1}(1+\pi_t)\frac{x_{1,t}}{x_{2,t}}, \tag{23}$$

where π_t^* is the reset price inflation rate.

Moreover, we can rewrite the price index definition in terms of the inflation rate as

$$(1+\pi_t)^{1-\varepsilon} = (1-\phi)(1+\pi_t^*)^{1-\varepsilon} + \phi. \tag{24}$$