Inflation - who cares?
Monetary Policy in Times of Low Attention

Oliver Pfäuti[*]

October 23, 2023

Link to most recent version

Abstract

I propose an approach to quantify *attention to inflation* in the data and show that the decrease in the volatility and persistence of U.S. inflation after the Great Inflation period was accompanied by a decline in the public's attention to inflation. This decline in attention has important implications (positive and normative) for monetary policy as it renders managing inflation expectations more difficult and can lead to *inflation-attention traps*: prolonged periods of a binding lower bound and low inflation due to slowly-adjusting inflation expectations. As attention declines the optimal policy response is to increase the inflation target. Accounting for the lower bound fundamentally changes the normative implications of declining attention. While lower attention raises welfare absent the lower-bound constraint, it decreases welfare when accounting for the lower bound.

JEL Codes: E31, E52, E58, E71
Keywords: Monetary Policy, Limited Attention, Inflation Expectations, Inflation Target, Effective Lower Bound

1 Introduction

Managing inflation expectations is an important instrument for monetary policy. It offers a powerful substitute to conventional tools if the nominal interest rate is constrained by the lower bound. By making promises about the future conduct of its policy, the monetary authority can shape inflation expectations, and thus, steer the real interest rate even when the nominal rate is constrained. At least this is how it works in theory.[1] But while traditional analyses assume that agents are perfectly informed and have rational expectations, recent empirical evidence suggests that the general public is usually poorly informed about and inattentive to monetary policy and inflation.[2] What do these low levels of attention imply for the conduct of monetary policy? And how has the public's attention to inflation changed over the last fifty years?

To tackle these questions, I propose an approach to quantify *attention to inflation* in the data. This approach is based on a model of optimal attention choice subject to information acquisition costs. The result is a law of motion for inflation expectations in which attention governs how strongly agents update their expectations following an inflation surprise. The optimal degree of attention depends positively on how volatile and persistent inflation is.

Using micro survey data of professional forecasters and consumers in the U.S., I then use this approach to estimate attention to inflation in the data and show that attention was very low in the years just before the Covid-19 crisis. In the 1970s and 1980s, on the other hand, attention to inflation was substantially higher. Consistent with the underlying model, times of higher inflation volatility and persistence are characterized by higher attention to inflation.

How does this decline in attention matter for the conduct of optimal monetary policy? To answer this question, I solve for the Ramsey optimal monetary policy in a standard New Keynesian model augmented with an effective lower bound (ELB) on the nominal interest rate and with inflation expectations that are characterized by limited attention. Both of these ingredients matter greatly for optimal monetary policy and the normative implications of declining attention. Lower attention has a stabilizing effect on inflation expectations and actual inflation, resembling more anchored (short-run) expectations. These stabilization benefits imply that lower attention is welfare improving—if we do not account for the ELB (similar to the findings in Paciello and Wiederholt (2014)). Accounting for the occasionally-binding ELB, however, completely overturns the normative implications of lower attention. Lower levels of attention

---

[1] For optimal monetary policy under full-information rational expectations, see e.g., Clarida et al. (1999), Woodford (2003) or Galí (2015). See Eggertsson and Woodford (2003), Adam and Billi (2006) or Coibion et al. (2012) for early analyses including the lower bound on nominal interest rates.

[2] Candia et al. (2021) and Coibion et al. (2023) show that U.S. firms as well as households are usually poorly informed about and quite inattentive to monetary policy. Coibion and Gorodnichenko (2012, 2015a), for example, show that models of limited attention more closely align with empirical patterns of inflation expectations, compared to models with full-information rational expectations. In line with limited attention, D'Acunto et al. (2022) show that forward guidance is quite ineffective in stimulating inflation expectations.

lead to a decline in welfare when we account for the lower-bound constraint. The reason is that lower attention renders managing expectations more difficult, which is particularly relevant if the nominal interest rate is constrained by the lower bound.[3]

To better understand these results, I first show that under sub-optimal policy, i.e., if monetary policy follows an *ad-hoc* Taylor rule, limited attention can lead to substantially longer periods at the ELB. Even though lower levels of attention attenuate the initial response of inflation expectations to a given shock, the decline in expectations becomes more persistent which hinders actual inflation from recovering quickly. Due to the persistently-low inflation, the monetary authority keeps the interest rate at the ELB for longer. I refer to these periods of long spells at the ELB and persistent declines in inflation and inflation expectations as *inflation-attention traps*. The response of the output gap, on the other hand, is very similar to the one under rational expectations. Thus, low attention offers a potential explanation for why inflation was relatively stable during the Great Recession but was persistently low during the subsequent recovery, seemingly disconnected from output (as documented in Del Negro et al. (2020)).

When replacing the *ad hoc* Taylor rule with the Ramsey optimal policy, I find that it is optimal to induce a higher average inflation rate to deal with these attention traps and the overall inability to manage inflation expectations at low levels of attention. With a higher average inflation rate, the nominal interest is also higher on average. This provides additional space when cutting the interest rate following adverse shocks, thus, mitigates the drawbacks of lower attention, and therefore helps to prevent long spells at the ELB.[4] Such an increase in the average inflation rate is not optimal if we abstract from the ELB, as then the nominal rate can go negative. Given the estimates of attention just before the Covid-19 crisis, the average inflation rate under Ramsey optimal policy is about 2-3 percentage points higher than under rational expectations. This increase in the level of inflation, however, is costly from a welfare perspective and it turns out that this level effect dominates the stabilization benefits. Thus, lower attention is welfare deteriorating when accounting for the ELB.

Another instrument to mitigate the drawbacks of limited attention are negative interest rate policies. Allowing for negative interest rates up to -0.5% (annualized) lowers the necessary increase in the optimal inflation target. As attention declines, however, the effectiveness of negative interest rate policies decreases and the optimal increase in the inflation target is close

---

[3]Coibion et al. (2022) and Coibion et al. (2023) show that managing expectations by the central bank is indeed a difficult task and the effects of monetary policy are much smaller than in most theoretical models. D'Acunto et al. (2022) find small effects of forward guidance on inflation expectations and durable consumption.

[4]Thus, the reason for the higher inflation rate is different from earlier papers that also consider an occasionally-binding ELB (e.g., Adam and Billi (2006), Adam et al. (2022)). In these papers, the higher average inflation rate arises due to promises the policymaker makes *at the lower bound*. In the present paper, on the other hand, the higher inflation rate arises due to considerations *before the lower bound* binds, *foreseeing what will happen* at the lower bound.

to the one without negative rates.

Related literature.   Bracha and Tang (2023) show that attention to inflation increases, when inflation increases. Using Google search data, Korenok et al. (2022) find that people's attention to inflation increases with inflation only after inflation exceeds a threshold of around 2-4%. Cavallo et al. (2017) and Weber et al. (2023) use randomized information treatments and show that attention of households and firms is higher in times of high inflation. Link et al. (2023) also find that attention to inflation is higher in times of high and volatile inflation. My key contribution relative to these papers is that I provide estimates of attention in a way that directly maps into otherwise standard macroeconomic models.

The empirical part of this paper is related to recent findings in Jørgensen and Lansing (2023) who show that inflation expectations have become more anchored over the last decades.[5] My measure of attention is inversely related to their definition of anchoring, but attention is concerned with short-run expectations whereas anchoring usually refers to the stabilization of long-run expectations. I complement their empirical analysis along several dimensions which I detail more closely in the empirical Section 2. Additionally, I offer new insights in how stabilized expectations matter when nominal interest rates are constrained by a lower-bound constraint and what this implies for optimal monetary policy.

My limited-attention model of inflation expectations is closely related to the general information choice problem in Maćkowiak et al. (2023). In contrast to their model, and the rational inattention literature more generally, agents in my model have a perceived law of motion of inflation that can potentially differ from the actual law of motion.[6] The reduced-form of the model that I bring to the data is close to the one in Vellekoop and Wiederholt (2019). In contrast to their paper, I focus on how attention changed over the last fifty years and show that attention tends to be higher in times of volatile and persistent inflation.

Ball et al. (2005), Adam (2007), Paciello and Wiederholt (2014) and Gáti (2022) characterize optimal monetary policy in models with different forms of limited attention. I contribute to this literature by allowing for an occasionally-binding lower-bound constraint. I show that accounting for the lower bound can lead to *qualitatively* different welfare implications due to changes in the optimal *level* of inflation. Wiederholt (2015) and Gabaix (2020) examine how information rigidities and inattention matter at the zero lower bound. Angeletos and Lian (2018) study the implications of relaxing the common knowledge assumption for forward guidance and show that the effects of forward guidance are attenuated in such a setting. My paper comple-

---

[5]Similarly, Gáti (2022) documents that anchoring of long-run inflation expectations is time varying and has substantially increased recently.

[6]Maćkowiak et al. (2023) provide a recent overview of this literature, which was inspired by the seminal paper Sims (2003). For further developments in this literature, see, among others, Mackowiak and Wiederholt (2009), Paciello and Wiederholt (2014), Maćkowiak et al. (2018), Afrouzi and Yang (2021), and see Gabaix (2019) for an overview of *behavioral* inattention.

ments these three papers by studying the Ramsey optimal policy in a fully stochastic setup and focuses on the implications for the optimal inflation target. To the best of my knowledge, the present paper is the first to study the trade off of lower attention in a fully stochastic model with an occasionally binding lower-bound constraint and to characterize the Ramsey optimal monetary policy in such a setting.

Outline. The rest of the paper is structured as follows. The empirical strategy to quantify attention, the description of the data and the empirical results are presented in Section 2. In Section 3, I show how limited attention can lead to inflation-attention traps, before I then study optimal policy in Section 4. Section 5 concludes. An online appendix provides all the derivations, several extensions and robustness checks.

2 Quantifying Attention

In this section, I derive an expectations-formation process under limited attention that provides a straightforward approach to measure attention to inflation empirically. The model is an application of Maćkowiak et al. (2023), who study a general problem of optimal information acquisition. I relegate all the details and derivations to appendix A.

The main difference to Maćkowiak et al. (2023) is that agents in my model do not exactly know the underlying process of inflation but have a simplified view of how inflation evolves.[7] In particular, the agent believes that (demeaned) inflation tomorrow, π', depends on (demeaned) inflation today, π, as follows

$$\pi' = \rho_\pi \pi + \nu,$$

where $\rho_\pi \in [0, 1]$ denotes the perceived persistence of inflation and $\nu \sim i.i.N.(0, \sigma_\nu^2)$. This assumption is supported by empirical evidence (see e.g., Faust and Wright (2013) or Canova (2007)).[8] Note, that the perceived volatility and persistence do not need to be the same as their actual counterparts, consistent with the empirical evidence on inflation expectations (see Table 7 in the Appendix). The agent wants to minimize her expected forecast error but inflation in the current period is unobservable and acquiring and processing information is costly. The agent thus faces a trade off how attentive she wants to be. I follow the literature on rational inattention and assume that the loss arising from making mistakes in her forecasts is quadratic with a scaling factor r, and the cost of information acquisition and processing is linear in mutual information with a scaling parameter λ.

---

[7]In the monetary model, later on, I will disentangle the effects from this potentially misperceived law of motion from the assumption of costly attention and show that most of the results are driven by limited attention and not the misperceived law of motion (see section 4.3.5).

[8]Fulton and Hubrich (2021) show that simple models such as AR(1) models are hard to beat when forecasting inflation in real time.

In this setup and with a normal prior, the optimal signal takes the form

$$s = \pi + \varepsilon,$$

with $\varepsilon \sim i.i.N.(0, \sigma_\varepsilon^2)$ (see Matějka and McKay (2015)).

The optimal forecast is given by $\pi^e = \rho_\pi E\left[\pi|s\right]$, and Bayesian updating implies

$$\pi^e = \rho_\pi (1 - \gamma)\hat{\pi} + \rho_\pi \gamma s, \tag{1}$$

where $\gamma = 1 - \frac{\sigma_{\pi|s}^2}{\sigma_\pi^2} \in [0,1]$ measures how much attention the agent pays to inflation, and $\hat{\pi}$ denotes the prior mean of π.

Solving for the optimal γ and writing the cost of information relative to the stakes, $\tilde{\lambda} \equiv \frac{\lambda}{r}$, yields the *optimal* level of attention, summarized in the following Lemma.

Lemma 1. *The optimal level of attention is given by*

$$\gamma = max\left(0, 1 - \frac{\tilde{\lambda}}{2\rho_\pi^2 \sigma_\pi^2}\right), \tag{2}$$

which shows that the optimal level of attention γ is

(i) decreasing in the relative cost of information acquisition, $\tilde{\lambda} \equiv \frac{\lambda}{r}$,

(ii) increasing in inflation volatility, σ_π, and

(iii) increasing in inflation persistence, ρ_π.

From Lemma 1, we see that aside from the relative information cost, $\tilde{\lambda}$, the persistence, ρ_π, and the volatility of inflation, σ_π, are crucial drivers of attention. The model predicts a positive relationship between attention and σ_π, as well as between attention and ρ_π. In the following, I will first estimate attention γ, asses how it has changed over time and then test whether there is indeed evidence for these positive relations.

2.1 Bringing the Model to the Data

To estimate attention in the data, I extend the law of motion of inflation expectations, equation (1), to a dynamic setup. The agent believes that inflation π follows

$$\pi_t = (1 - \rho_\pi)\bar{\pi} + \rho_\pi \pi_{t-1} + \nu_t,$$

where $\bar{\pi}$ is the agent's long-run belief about inflation and ρ_π is the perceived persistence of inflation. I assume that the error term ν_t is normally distributed with mean zero and variance

6

σ_ν^2.[9]

The agent receives a signal about inflation of the form

$$s_{it} = \pi_t + \varepsilon_{it},$$

where the noise ε_{it} is assumed to be normally distributed with variance σ_ε^2.

Given these assumptions, it follows from the (steady state) Kalman filter that optimal updating is given by

$$\pi_{t+1|t,i}^e = (1 - \rho_\pi)\bar{\pi} + \rho_\pi \pi_{t|t-1,i}^e + \rho_\pi \gamma \left(\pi_t - \pi_{t|t-1,i}^e \right) + u_{i,t}, \tag{3}$$

where the updating gain γ captures the agent's level of attention, and $u_{i,t} = \rho_\pi \gamma \varepsilon_{i,t}$ is a scaled version of the i.i.d. noise term $\varepsilon_{i,t}$. From equation (3), we observe that lower attention, i.e., a lower γ, implies that the agent updates her expectations to a given forecast error, $(\pi_t - \pi_{t|t-1,i}^e)$, less strongly. Lower attention is reflected in more noisy signals, and more noise means the agent trusts her received signals less and thus, puts less weight on these signals. Hence, her expectations remain more strongly anchored at her prior beliefs.

In the estimation of equation (3), I allow for individual-specific intercepts. This can either reflect a mean bias in the perceived inflation rate, $\bar{\pi}_i \neq \bar{\pi}$, or that the agent believes her signals are biased on average, as in Vellekoop and Wiederholt (2019).

2.2 Data

I focus on the Survey of Professional Forecasters (SPF) from the Federal Reserve Bank of Philadelphia, as well as the Survey of Consumers from the University of Michigan (SoC). In the Appendix, I show that the findings extend to other data sets as well. For the SPF, I consider individual and aggregate forecasts. The main focus is on expectations about the quarter-on-quarter percentage change in the GDP deflator, which is available since 1969. I drop forecasters for which I have less than eight observations. As a robustness check, I will show that the results are robust to using expectations about the consumer price index, CPI. This data series, however, is only available since 1979.

While the SPF provides data on expectations about the next quarter, the SoC only provides one-year-ahead expectations. Therefore, I will compare them to the actual year-on-year changes in the CPI.[10] As the SoC does not have a panel dimension, I consider average (and median) expectations. Additionally, I estimate attention using the Survey of Consumer Expectations from the Federal Reserve Bank of New York (SCE). The SCE, launched in 2013, has a panel

---

[9]In Appendix B.1, I discuss the case in which the agent believes that inflation follows an AR(2) process.

[10]The question in the SoC is not explicitly about the CPI but about "prices". Using GDP deflator inflation instead of CPI inflation barely affects the results.

structure and thus allows me to estimate attention using individual-consumer data, at least for the period after 2013. Data on actual inflation comes from the FRED database from the Federal Reserve Bank of St. Louis. Throughout, I focus on the pre-Covid period and end the sample in 2019Q4.[11] Appendix B provides summary statistics and plots the discussed time series.

Jørgensen and Lansing (2023) estimate how strongly anchored inflation expectations in the U.S. are and how this changed over the last fifty years. I extend their empirical strategy in several dimensions. First, I allow the persistence of perceived inflation to change over time and do not restrict it to follow a random walk. Second, I show that not only aggregate professional forecasters' expectations have become more anchored, but also consider individual-specific expectations, as well as consumers' inflation expectations. Third, I do not impose the structure of the New Keynesian Phillips Curve on the data but directly estimate attention simply based on the proposed law of motion for inflation expectations. That said, we will see that my results are consistent with the results in Jørgensen and Lansing (2023).

2.3 Estimation Results

Before estimating attention, I rewrite the updating equation (3) as

$$\pi^e_{t+1|t,i} = \beta_i + \beta_1 \pi^e_{t|t-1,i} + \beta_2 \left(\pi_t - \pi^e_{t|t-1,i} \right) + u_{i,t}, \tag{4}$$

where $\beta_i = (1 - \rho_\pi)\bar{\pi}_i$, $\beta_1 = \rho_\pi$ and $\frac{\beta_2}{\beta_1} = \gamma$. For the SPF, where I can use individual forecasts, I estimate (4) using a forecaster-fixed-effects regression. Since the dependent variable shows up with a lag on the right-hand side, however, a standard fixed-effects regression introduces a bias (Nickell, 1981). Therefore, I apply the estimator proposed by Blundell and Bond (1998) (BB for short) and I use all available lags of the dependent variable as instruments.[12] All reported standard errors are robust with respect to heteroskedasticity and serial correlation. As an alternative to the BB estimator, I also estimate (4) using pooled OLS. For the Survey of Consumers I cannot include individual fixed effects as I use average and median expectations and thus, I apply the Newey-West estimator using four lags (Newey and West (1987)).

To examine how attention changed over time, I run regression (4) for the period before and after 1990, separately. The results are robust to different split points (see Appendix B.1). Later on, I will estimate (4) using rolling-windows of ten years each and show that the general patterns I document are robust.

Table 1 shows the results. Here, $\widehat{\gamma}_{pre}$ and $\widehat{\gamma}_{post}$ denote the estimated attention parameters

---

[11]In a follow-up paper, I study the implications of changes in people's attention to inflation in driving high and persistent inflation and show that this offers a potential explanation for the inflation surge that occurred after the Covid-19 pandemic (Pfäuti, 2023).

[12]Appendix B.1 shows that the results are robust to using fewer lags.

Table 1: Regression Results of Equation (4)

	Professional Forecasters		Consumers	
	Blundell Bond	Pooled OLS	Averages	Median
$\widehat{\gamma}_{pre}$	0.70	0.44	0.75	0.43
s.e.	(0.1005)	(0.0397)	(0.1574)	(0.0970)
$\widehat{\gamma}_{post}$	0.41	0.22	0.31	0.24
s.e.	(0.0522)	(0.0290)	(0.0881)	(0.0601)
N	3566	3566	120	120

Note: This table shows the results from regression (4) for professional forecasters (SPF) as well as for consumers. For the SPF, I use the Blundell and Bond (1998) (BB) estimator (first two columns), as well as pooled OLS (columns 3-4). For the Survey of Consumer, I consider average expectations (columns 5-6) and median expectations (columns 7-8). $\widehat{\gamma}_{pre}$ and $\widehat{\gamma}_{post}$ denote the estimated attention parameters for the period pre 1990 and post 1990, respectively. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

for the period pre 1990 and post 1990, respectively. We see that attention is substantially lower after 1990 compared to the period before 1990.[13] This is true for professional forecasters and for consumers, and as I show in Appendix B.1, also for other datasets. The point estimates after 1990 are basically half of what they were before the 1990s and these differences are statistically significant at the 1% level.

The decline in attention is even more pronounced when focusing on the most recent decade. To show this, I run regression (4) for the period between 2010 and 2020. Additionally, I also use data from the New York Fed Survey of Consumer Expectations, starting in 2013. The advantage of this survey compared to the Michigan Survey is that it surveys the same consumers up to twelve times in a row, providing a much larger sample size and a panel dimension. Table 2 shows the results. We see that overall, attention declined substantially compared to earlier periods and is between 0.04 and 0.17 during this period of low and stable inflation. Furthermore, the results from the SCE lie in the same ballpark as the ones from the Michigan Survey, which indicates that using average (or median) consumer expectations does not fundamentally affect the results. In fact, the estimated attention parameter for the average expectations from the Michigan Survey is 0.04 when restricting the sample to 2013-2020, which is exactly the same as the estimate obtained from the New York Fed Survey.

2.3.1 When is Attention High?

Lemma 1 states that two key drivers of attention are inflation volatility and inflation persistence. To examine these relationships empirically, I estimate regression (4), using a rolling-window

---

[13]I test the validity of the instruments in the Blundell-Bond estimation by testing for autocorrelation of order one and two in the first-differenced error terms. The respective p-values are 0.000 (order 1) and 0.973 (order 2) for the period before 1990 and 0.000 (order 1) and 0.737 (order 2) for the period after 1990. This indicates that the instruments used in the estimation are valid.

Table 2: Attention since 2010

| | Professional Forecasters | | Consumers | | NY Fed Survey |
	Blundell Bond	Pooled OLS	Averages	Median	Pooled OLS
$\widehat{\gamma}$	0.17	0.07	0.12	0.09	0.04
s.e.	(0.0729)	(0.0333)	(0.0658)	(0.0616)	(0.0316)
N	1322	1322	40	40	74229

Note: This table shows the results from regression (4) for the period between 2010 and 2020 for professional forecasters (SPF) as well as for consumers. For the SPF, I use the Blundell and Bond (1998) (BB) estimator (first column), as well as pooled OLS (column 2). For the Survey of Consumer, I consider average expectations (column 3) and median expectations (columns 4). The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation. Additonally, column 5 shows the results for consumer inflation expectations from the New York Fed Survey of Consumer Expectations.

approach in which every window is 10 years long, and estimate one attention parameter for every window, $\widehat{\gamma}_t$, as well as the period-specific inflation volatility, $\widehat{\sigma}_{\pi,t}$ and the persistence parameter, $\widehat{\rho}_{\pi,t}$. In particular, I use the window-specific standard deviation of inflation as my measure of $\widehat{\sigma}_{\pi,t}$ and the first-order autocorrelation of inflation for $\widehat{\rho}_{\pi,t}$. Appendix B.1 shows that the following results also hold for different window lengths or when using the standard deviation and persistence of *expected* inflation for $\widehat{\sigma}_{\pi,t}$ and $\widehat{\rho}_{\pi,t}$.

Figure 1 summarizes the results graphically. The left scatterplot shows the inflation volatility on the horizontal axis and the estimated attention parameter, $\widehat{\gamma}$, on the vertical axis. The attention parameters shown in the figure are the ones for individual professional forecasters, obtained via pooled OLS. The right panel shows the relationship between attention and inflation persistence (on the horizontal axis). In both cases, we see that there is a clear positive relationship, just as the limited-attention model predicts.

To check if these findings are statistically significant, I regress attention on inflation volatility or on inflation persistence as follows

$$\widehat{\gamma}_t = \alpha_1 + \beta\widehat{\sigma}_{\pi,t} + u_t \tag{5}$$

$$\widehat{\gamma}_t = \alpha_2 + \zeta\widehat{\rho}_{\pi,t} + v_t. \tag{6}$$

Table 3 reports the results. Standard errors are robust with respect to heteroskedasticity. We see that the observed patterns in Figure 1 are indeed statistically significant. Attention to inflation is positively correlated with inflation volatility and inflation persistence, as Lemma 1 predicts. This is true for professional forecasters as well as for consumers. Overall, the magnitudes of the estimates indicate that the results are somewhat stronger for professional forecasters. Appendix B.1 shows that these results hold when regressing attention on inflation volatility and persistence jointly. I also show that the results are robust when controlling for the *average level* of inflation and the average level of inflation has no significantly-positive

Figure 1: Attention, Inflation Volatility and Inflation Persistence

(a) Inflation Volatility

(b) Inflation Persistence
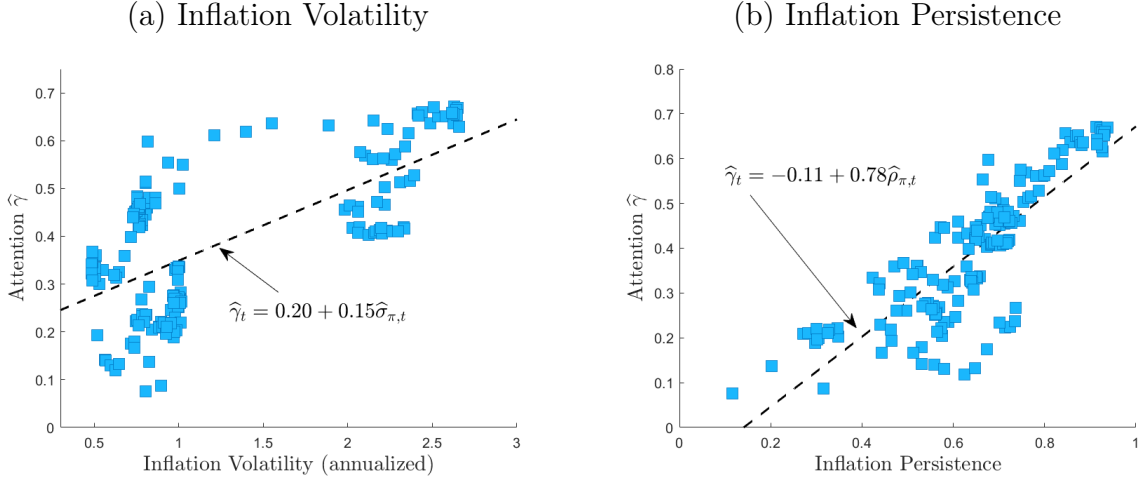
Notes: The right panel shows the relationship between inflation volatility and the estimated attention parameter, $\widehat{\gamma}$. The left panel shows the relationship between the persistence of inflation and the estimated attention parameter. Both panels report the results for individual professional forecasters where the attention parameter was estimated via pooled OLS.

Table 3: Attention, Inflation Volatility and Inflation Persistence

	Professional Forecasters		Consumers
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.13***	0.15***	0.09***
s.e.	(0.0155)	(0.0105)	(0.0272)
$\widehat{\zeta}$	0.71***	0.78***	0.56***
s.e.	(0.0568)	(0.0349)	(0.0714)
N	165	165	163

Note: This table shows the results of regression (5) and (6). Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

effect on attention when controlling for volatility and persistence. Given the decline in inflation volatility and inflation persistence over the last fifty years (see Table 7 in Appendix B), the positive correlation with attention supports the findings in Table 1: attention declined as the volatility and persistence of inflation declined.[14]

Additional evidence and robustness. In Appendix B.1, I show that the presented results are robust to using different data sources, different sample splits, different specifications of the BB estimator, allowing for time-fixed effects to account, for example, for varying trend inflation, as well as constructing a quasi panel of consumers, based on their income. I further discuss the case in which the agent believes that inflation follows an AR(2) process. I again find

---

[14]Benati (2008) documents a decline in inflation persistence in advanced economies, especially for countries that introduced inflation targeting regimes.

that attention has declined substantially after the 1990s and that the additional coefficients showing up due to the AR(2) rather than the AR(1) tend to be insignificant. In Appendix B.2, I document a decline in news coverage of inflation in popular news papers as well as in books, thus, providing additional, complementary evidence on the decline in attention after the 1980s.[15] Additionally, I show that another measure of attention (based on survey respondents answering "I don't know" when asked about their inflation expectations) is strongly correlated with my proposed measure of attention. Finally, I provide additional evidence favoring the proposed attention model compared to a setup in which the agent cannot distinguish between a trend and a cyclical component with time-varying volatilities of these two components. In that case, if the trend component's contribution to overall inflation increases, the agent's forecast would become more responsive to current inflation, too. I show that professional forecasters' *nowcasts of inflation* are more accurate in times of high inflation volatility and persistence, which supports the proposed attention model rather than this alternative model.

3   Monetary Policy Implications of Limited Attention

How does the decline in attention to inflation affect the conduct of monetary policy? To answer this question, I augment the standard New Keynesian model with inflation expectations that are characterized by limited attention, and a lower-bound constraint on the nominal interest rate (model details and all derivations are in Appendix C). I build on the standard New Keynesian model without capital, with rigid prices in the spirit of Rotemberg (1982) and with a lower bound on the nominal interest rate. The government pays a subsidy to intermediate-goods producers to eliminate steady state distortions arising from market power. I focus on the case with zero-steady state inflation as my baseline case, but discuss the case of positive trend inflation in Section 4.3.3.

The linearized model consists of an aggregate supply equation, the *New Keynesian Phillips Curve*, and an aggregate Euler (or IS) equation:

$$\pi_t = \beta \pi_{t+1|t}^e + \kappa y_t^{gap} + u_t, \tag{7}$$

$$y_t^{gap} = E_t y_{t+1}^{gap} - \varphi \left(i_t - \pi_{t+1|t}^e - r_t^n \right), \tag{8}$$

where κ measures the sensitivity of aggregate inflation to changes in the output gap, y_t^{gap}, $\beta \in (0,1)$ denotes the time discount factor of the representative household, and u_t are cost-push shocks, following an AR(1) process with persistence $\rho_u \in [0,1]$ and innovations $\varepsilon^u \sim i.i.N.(0, \sigma_u^2)$. The output gap is the log deviation of output from its efficient counterpart

---

[15] Carroll (2003) proposes a micro-foundation of sticky information models (as in Mankiw and Reis (2002)) that relies on news coverage of inflation. Lamla and Lein (2014) and Pfajfar and Santoro (2013) test these predictions empirically.

that would prevail under flexible prices. Altogether, equation (7) summarizes the aggregate supply side of the economy. Equation (8), together with monetary policy, determines aggregate demand in this model. Here, $\varphi > 0$ measures the real rate elasticity of output, i_t is the nominal interest rate which is set by the monetary authority, and r_t^n is the natural interest rate. The natural interest rate is the real rate that prevails in the economy with fully flexible prices and is exogenous. It follows an AR(1) process with persistence $\rho_r \in [0,1]$ and innovations $\varepsilon^r \sim i.i.N.(0, \sigma_r^2)$, independent of ε^u. The nominal interest rate and the natural rate are both expressed in absolute deviations of their respective steady state values, $\bar{i}$ and $\bar{r}^n$. E_t denotes the full-information rational expectations operator. I relax the assumption that agents have rational expectations about the output gap in section 4.3 and in appendix E.1.

Inflation expectations are characterized by limited attention and are given by

$$\pi_{t+1|t}^e = (1 - \rho_\pi)\bar{\pi} + \rho_\pi \pi_{t|t-1}^e + \rho_\pi \gamma \left(\pi_t - \pi_{t|t-1}^e\right), \tag{9}$$

where the notation is the same as in Section 2. Given the representative agent assumption, I abstract from noise shocks in (9). The law of motion of inflation expectations, equation (9), is driven by two main assumptions. First, the perceived law of motion follows an AR(1) process and second, paying attention is costly. In Section 4.3.5, I will disentangle the two and show that it is mainly the second assumption that drives the implications for optimal monetary policy. For the most part, I will focus on $\rho_\pi = 1$ in which case average inflation expectations align with actual average inflation and long-run beliefs $\bar{\pi}$ are irrelevant. I discuss the case with $\rho_\pi < 1$ in Appendix E.4. For empirically-realistic values of ρ_π the results are very similar to the case with $\rho_\pi = 1$. This belief formation process is empirically plausible, in the sense that it is consistent with recent empirical findings, documented in Angeletos et al. (2021): after a shock, expectations initially underreact, followed by a delayed overreaction (see Appendix E.3).

As is standard in the rational inattention literature, I assume that the attention parameter γ is constant.[16] The usual assumption to obtain this is that in period $t = 0$ the agent chooses her level of attention and then obtains all future signals at this point. This leaves conditional second moments time-invariant and thus, the optimal level of attention constant. I will, however, compare economies with different levels of attention. In Section 4.3.4, I discuss the case in which attention to inflation is time varying.

Calibration. I calibrate the model to quarterly frequency. I assume an annualized steady state natural rate of 1%. The rest of the calibration is taken from Adam and Billi (2006). Table 4 summarizes the calibration. The attention parameter γ will be varied to understand its role for monetary policy.

---

[16]See, e.g., Mackowiak and Wiederholt (2009), Maćkowiak et al. (2018, 2023).

Table 4: Model Parameterization

Parameter	Value	Source/Target
Preferences and technology		
β	0.9975	Average natural rate of 1%
φ	1	Adam and Billi (2006)
κ	0.057	Adam and Billi (2006)
Exogenous shock processes		
ρ_r	0.8	Adam and Billi (2006)
σ_r	0.2940%	Adam and Billi (2006)
ρ_u	0	Adam and Billi (2006)
σ_u	0.154%	Adam and Billi (2006)

Attention and the Phillips Curve. Inflation expectations are a crucial driver of actual inflation through the supply side of the economy, and thus, changing levels of attention affect the Phillips Curve. The following Proposition summarizes these effects.

Proposition 1. *The New Keynesian Phillips Curve under limited attention is given by*

$$\pi_t = \frac{\beta(1-\gamma)}{1-\beta\gamma}\pi^e_{t|t-1} + \frac{\kappa}{1-\beta\gamma}y^{gap}_t + \frac{1}{1-\beta\gamma}u_t.$$

Proof. See Appendix D.2. □

From Proposition 1, we see that for a given prior inflation expectation, $\pi^e_{t|t-1}$, and a given realization of the exogenous shock u_t, inflation becomes less sensitive to changes in the output gap at lower levels of attention. Put differently, a decrease in attention γ resembles a flatter Phillips Curve. We also see that for a given output gap, inflation becomes less sensitive to cost-push shocks u_t at lower levels of attention.

Thus, Proposition 1 captures the stabilizing effects of lower attention. As attention declines, firms' inflation expectations react less to changes in actual inflation. Through the Phillips Curve, this muted reaction of expectations in turn stabilizes inflation itself. What cannot be seen from Proposition 1, but what will be crucial in the subsequent analysis, is that lower attention not only affects the initial response of inflation and inflation expectations but the dynamics as well. Changes in inflation and inflation expectations become more persistent at low levels of attention, even though the initial response is muted. I now show in a numerical example that at low levels of attention the economy can get stuck in an *inflation-attention trap*.[17]

---

[17]In Appendix D, I show analytically how lower levels of attention mute the effects of forward guidance.

3.1 Inflation-Attention Traps

To close the model from Section 3, I assume for now that, away from the lower bound, the monetary authority sets the nominal interest rate according to a Taylor rule

$$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1 - \rho_i)(\phi_\pi \pi_t + \phi_y y_t^{gap}), \tag{10}$$

where $\rho_i \in [0, 1)$ captures interest rate smoothing, $\phi_\pi > 1$ and $\phi_y \geq 0$ denote the reaction coefficients to inflation and the output gap, respectively. The actual interest rate, i_t, however is constrained by the lower bound

$$i_t = max\{\tilde{i}_t, -\bar{i}\}, \tag{11}$$

where I set the lower bound (in levels) to zero.

I set the persistence parameter of the nominal interest rate to 0.7, and the reaction coefficients $\phi_\pi = 2$ and $\phi_y = 0.5$, as in Andrade et al. (2019). In Appendix E.2, I show that the exact specification of the Taylor rule is inconsequential for the following results.
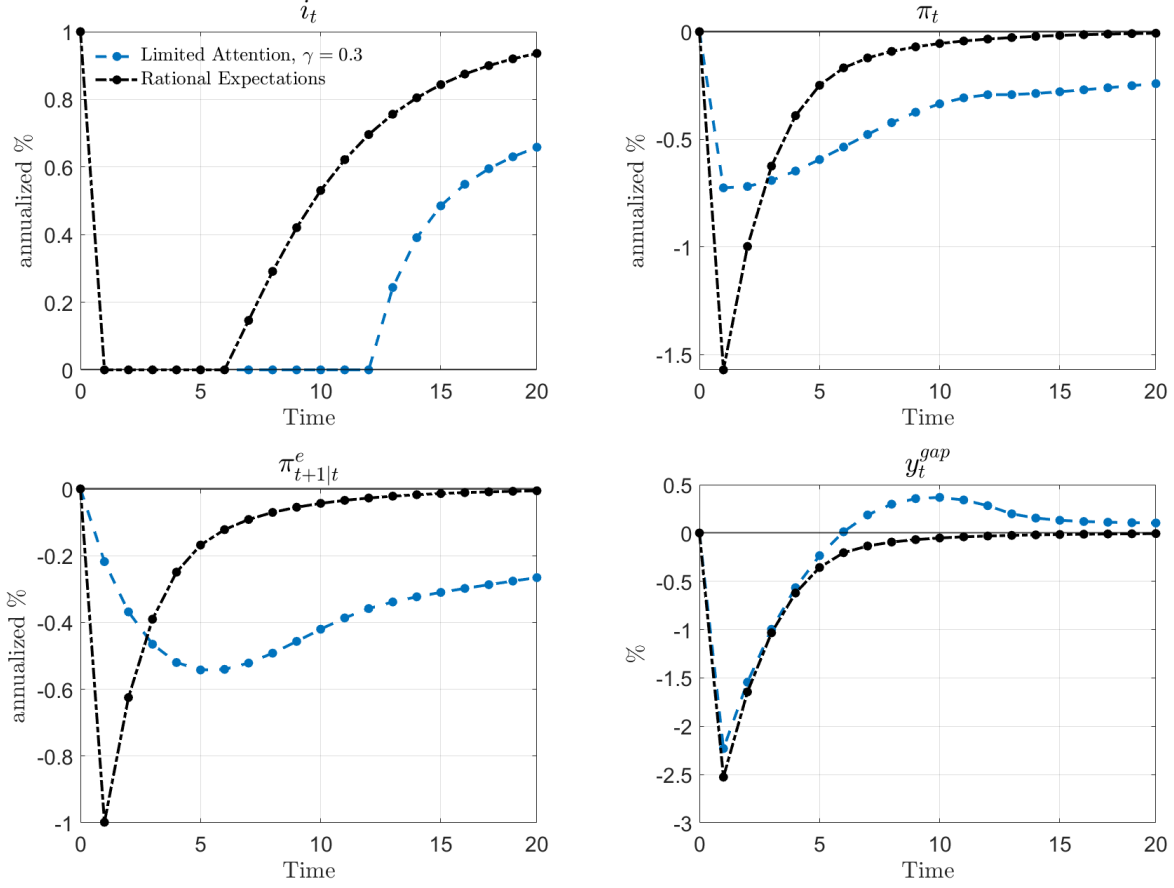
Figure 2 plots the impulse response functions of the model's main variables to a negative natural rate shock of three standard deviations that pushes the nominal interest rate to the lower bound. The black-dashed-dotted lines are the IRFs in the model under FIRE and the blue-dashed lines are the ones under limited attention for the case $\gamma = 0.3$.

In both cases, the shock is large enough to push the economy to the lower bound. While the reaction of the output gap is very similar in both economies, the responses of inflation and inflation expectations are strikingly different. Initially, the muted response of inflation expectations to the adverse shock is reflected in a smaller downturn of inflation itself under limited attention. This captures the stabilizing effects that come with lower attention. The sluggish adjustment of inflation expectations in the following, however, leads to a very persistent undershooting of inflation. Even five years after the shock, inflation and inflation expectations are still substantially below their steady state levels of zero. The result is a prolonged period of a binding lower bound. While the economy under rational expectations escapes the ELB six periods after the shock, the economy under limited attention is stuck for twice as long. This is what I label *inflation-attention trap*. A side-effect of these traps is that the long ELB period leads to an output boom. As discussed earlier, this (expected) output boom in the future, however, has rather small effects on the economy today if people are inattentive.

Overall, limited attention to inflation offers a possible explanation for why several advanced economies were stuck at the ELB after the financial crisis, as well as inflation that undershot the central banks' inflation targets, even though the initial decrease was muted and output recovered quite strongly after declining severely initially (Del Negro et al. (2020)). In other words, the limited-attention model can explain the *missing deflation puzzle* as well as the

Figure 2: Impulse Response Functions to a Negative Natural Rate Shock

Note: This figure shows the impulse-response functions of the nominal interest rate (upper-left panel), inflation (upper-right panel), inflation expectations (lower-left) and the output gap (lower-right) to a negative natural rate shock of three standard deviations. The blue-dashed lines show the case for the limited-attention model and the black-dashed-dotted lines for the rational expectations model. Everything is in terms of percentage deviations from the respective steady state levels, except the nominal rate is in levels.

missing inflation puzzle (Coibion and Gorodnichenko (2015b), Constancio (2015)).

4   Optimal Monetary Policy

To understand how monetary policy should optimally deal with declining attention, I now derive the Ramsey optimal monetary policy in this economy. I focus on the case of $\rho_\pi = 1$, in which average inflation expectations coincide with the actual inflation average. Appendix E.4 reports the results when relaxing this assumption.

The policymaker's objective is to maximize the representative household's utility, taking the household's and firms' optimal behavior, including their attention choice, as given. Thus, the policymaker cannot exploit the private agent's lack of information. Nevertheless, the policymaker can affect inflation expectations by influencing inflation itself and can set the average

16

inflation expectations by setting the average inflation rate.

The policymaker is paternalistic in the sense of Benigno and Paciello (2014) and evaluates the household's utility under rational expectations. A second-order approximation to the household's utility function yields the policymaker's objective

$$-\frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \chi \left(y_t^{gap}\right)^2\right], \tag{12}$$

where χ is the relative weight of the output gap, which I set to $\chi = 0.007$ as in Adam and Billi (2006). In the following, I refer to (12) as *welfare*.

In sum, the optimal policy problem is given by

$$\max_{\pi_t, y_t^{gap}, i_t} \quad -\frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \chi \left(y_t^{gap}\right)^2\right] \tag{13}$$

subject to

$$\pi_t = \beta \pi_{t+1|t}^e + \kappa y_t^{gap} + u_t \tag{14}$$

$$y_t^{gap} = E_t y_{t+1}^{gap} - \varphi \left(i_t - \pi_{t+1|t}^e - r_t^n\right) \tag{15}$$

$$\pi_{t+1|t}^e = \pi_{t|t-1}^e + \gamma \left(\pi_t - \pi_{t|t-1}^e\right) \tag{16}$$

$$u_t = \rho_u u_{t-1} + \varepsilon_t^u \tag{17}$$

$$r_t^n = \rho_r r_{t-1}^n + \varepsilon_t^r \tag{18}$$

$$i_t \geq -\bar{i}, \tag{19}$$

with $\varepsilon_t^u \sim i.i.N.\left(0, \sigma_u^2\right)$ and $\varepsilon_t^r \sim i.i.N.\left(0, \sigma_{r^n}^2\right)$ and (19) is the lower-bound constraint.[18] All variables are in percent deviations from their respective steady state, except the nominal interest rate and the natural rate which are in absolute deviations.

## 4.1 The Optimal Inflation Target

What do low levels of attention imply for inflation volatility and the optimal inflation target? For this, I solve the Ramsey problem for different levels of attention, namely $\gamma \in \{0.05, 0.1, 0.2, 0.3\}$. An attention parameter of 0.3 is close to the estimates for consumers' attention after 1990, and the lower levels of 0.05 and 0.1 are close to the ones observed since 2010.

Figure 3 shows the results. The average inflation rate under Ramsey optimal policy—what I refer to as the optimal inflation target—is plotted in the left panel and the inflation volatility in

---

[18]I solve this numerically by recursifying the constrained optimization problem, as in Marcet and Marimon (2019).

Figure 3: Optimal Inflation and Inflation Volatility
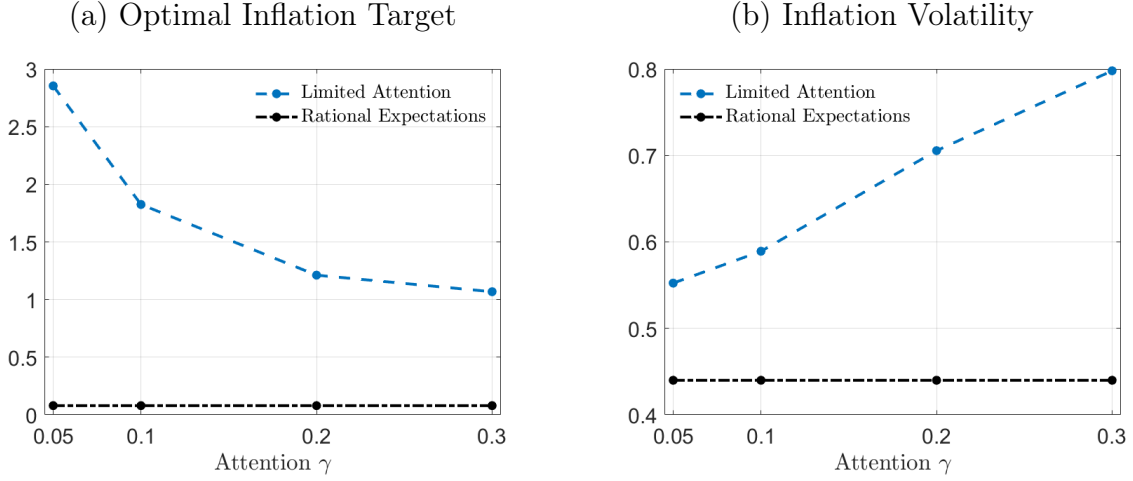
(a) Optimal Inflation Target   (b) Inflation Volatility

Notes: This figure shows the average inflation rate under Ramsey optimal policy (left panel) and inflation volatility for different attention levels. The blue-dashed lines show the results for the model under limited attention, and the black-dashed-dotted lines for the rational-expectations model.

the right panel. The blue-dashed lines show the results for the model under limited attention, and the black-dashed-dotted lines for the rational-expectations model. We see that the optimal inflation target increases substantially as attention declines. At the levels of attention estimated just before the Covid crisis, $\gamma \in \{0.05, 0.1\}$, the inflation target is about 2-3 percentage points higher than under rational expectations due to the discussed ineffectiveness of forward-guidance policies. By increasing the average level of inflation the average nominal interest rate increases and thus makes it less likely that the ELB becomes binding. Indeed, the frequency of a binding ELB decreases substantially. For $\gamma = 0.3$, the ELB is binding 22% of the time under optimal policy, whereas this value shrinks to 1.4% for an attention level of 0.05.

While lower attention renders forward guidance, make-up policies and other policies that work (partly) through inflation expectations less effective, lower attention also stabilizes inflation, as can be seen from the right panel in Figure 3. First, because lower attention mutes the inflation response to shocks and output (see Proposition 1). Second, the lower ELB frequency further stabilizes the economy. Thus, lower attention to inflation can help stabilizing actual inflation and reduces the number of binding-ELB periods. The lower inflation volatility at lower levels of attention in fact justifies these low attention levels, as optimal attention depends positively on inflation volatility (see Section 2). This low volatility, however, requires an increase in the inflation target, which is costly. Thus, it is not clear *a priori* whether lower attention leads to welfare gains or not.

4.2 Welfare

What are the effects of declining attention on overall welfare? Welfare is given by equation (12) and from the previous discussion, we know that lower attention poses a trade off. On the one hand, inflation volatility decreases and the ELB binds less frequently, when attention is low. This raises welfare. On the other hand, lower attention complicates managing inflation expectations and thus, the optimal average *level* of inflation increases, which is costly. Which effect dominates?

Panel (a) in Figure 4 shows that the cost of the level effect outweighs the stabilization benefits. As attention falls, welfare decreases. This is especially pronounced at low levels of attention, where the optimal inflation target increases substantially (see Figure 3).

Figure 4: Welfare and Attention

(a) ELB

(b) No ELB
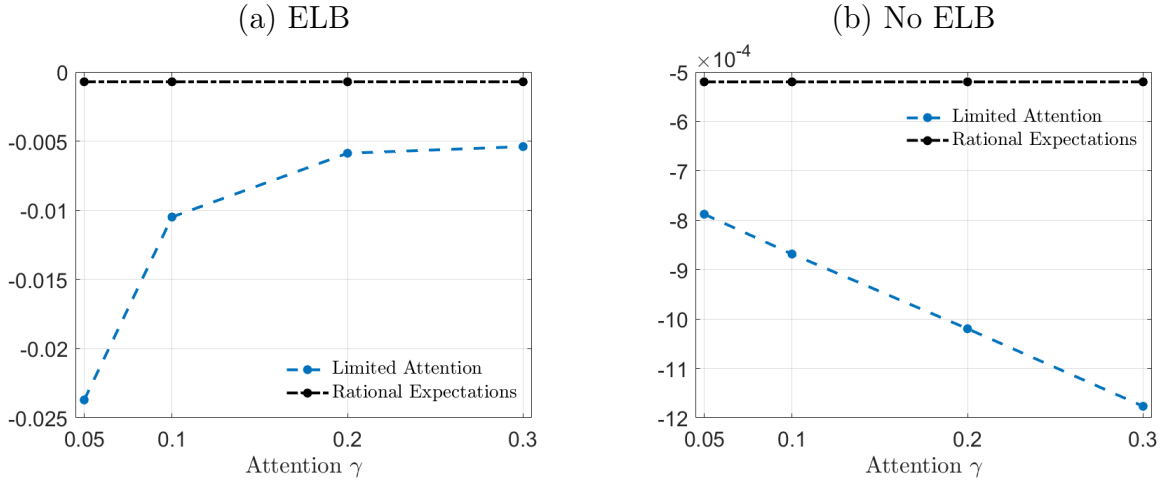
Notes: This figure shows welfare (12) under Ramsey optimal policy for different levels of attention. The left panel shows the results for the case with an occasionally-binding ELB, and the right panel without an ELB. The blue-dashed lines show the results for the model under limited attention, and the black-dashed-dotted lines for the rational-expectations model.

Absent the lower-bound constraint, the complications in managing inflation expectations due to limited attention are much less pronounced since managing expectations is particularly important at the lower bound. In fact, lower attention is welfare improving in the case without an ELB. Panel (b) in Figure 4 shows this graphically. The stabilization benefits that arise from lower attention—which is reflected in more anchored expectations—lead to an increase in welfare.

These findings show that accounting for the ELB is crucial for making a normative statement about costs and benefits of stabilizing inflation expectations. The ELB highlights the drawbacks that arise from the stabilization of expectations due to the fall in attention, as the management of expectations becomes particularly relevant when the ELB binds.

4.3 Extensions

In this section, I present several extensions of the baseline model and show that the overall implications derived so far remain robust, even when allowing (i) for a negative ELB, (ii) for non-rational output gap expectations, (iii) for a time-varying slope of the Phillips Curve or with non-zero trend inflation, and (iv) with time-varying attention to inflation. In 4.3.5, I further show that the results are largely driven by the agents' limited attention rather than by their perceived law of motion.

4.3.1 Negative Interest Rate Policies

In recent years, several central banks in advanced economies have implemented negative interest rate policies (NIRP).[19] Could negative rates limit the negative consequences of declining attention? In order to answer this question, I solve the same Ramsey optimal policy problem as above, but set the effective lower bound to -0.5% (annualized).

Figure 5: Negative Interest Rate Policies and Attention

(a) Optimal Inflation Target   (b) Change in Inflation Target
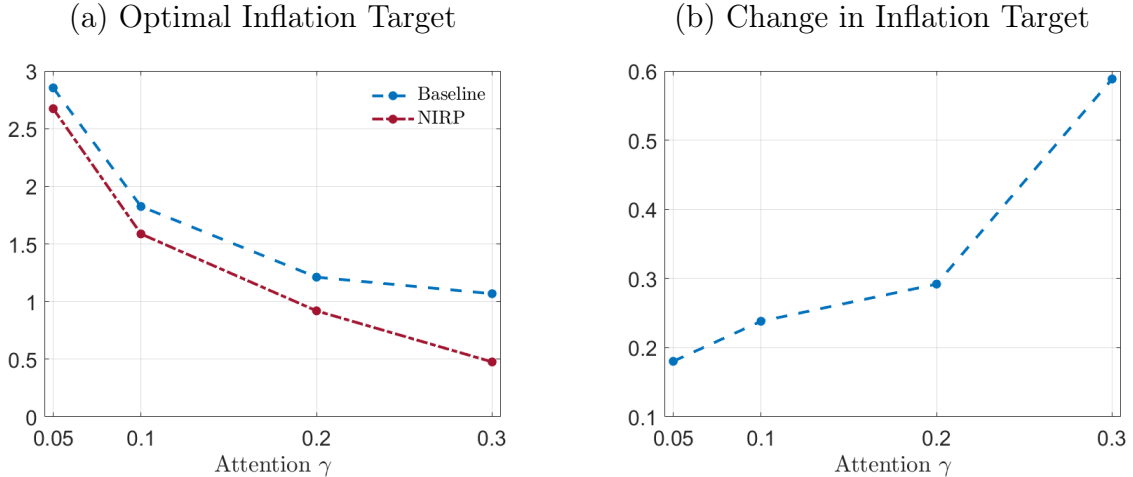
Notes: The left panel shows the average inflation rate under optimal policy for different degrees of attention γ. The blue-dashed lines show the results for the benchmark model where the lower bound is at 0, and the red-dashed-dotted lines show the results when allowing for negative interest rates up to -0.5% (annualized). The right panel shows the difference in the optimal inflation targets, defined as $\pi^{*,ZLB} - \pi^{*,NIRP}$, where π^* denotes the optimal inflation target and the superscripts ZLB and $NIRP$ denote the two cases where the ELB is at 0% or -0.5%, respectively.

Figure 5 reports the outcomes. Panel (a) shows the optimal inflation target (red-dashed-dotted line) and compares it to the case with an ELB at 0 (blue-dashed line). We see that the additional policy space due to the negative lower bound indeed calls for a lower inflation target. However, the decline in attention also weakens the effectiveness of NIRP. We see this by observing that the optimal inflation target under NIRP gets closer to the one without negative

---

[19]See Brandao-Marques et al. (2021) for a recent survey on negative interest rate policies and its effectiveness.

rates as attention declines. To see these gaps clearly, panel (b) shows the difference in the optimal inflation targets, defined as $\pi^{*,ZLB} - \pi^{*,NIRP}$, where π^* denotes the optimal inflation target and the superscripts ZLB and $NIRP$ denote the two cases where the ELB is at 0% or -0.5%, respectively. Overall, allowing for negative policy rates can help limiting the drawbacks of low attention but these policies itself become less effective as attention declines.

4.3.2   Non-Rational Output Gap Expectations

So far, I assumed that output gap expectations are fully rational. I now relax this assumption and show that in this case, (i) the optimal inflation target further increases and (ii) non-rational output gap expectations are welfare deteriorating. But before going into these results, I need to quantify people's attention to the output gap. As expectations about the output gap are not available, I use expectations about unemployment changes over the next year from the Michigan Survey and then estimate attention to unemployment γ^y as in Section 2. As I detail in Appendix E.1, I estimate attention to unemployment to have slightly increased from 0.09 to 0.1 from the period before 1990 to the period after 1990. This difference, however, is not statistically significant. Thus, I do not find any evidence for changes in people's attention to unemployment over the same period in which their attention to inflation decreased.

To understand the monetary policy implications of non-rational output gap expectations in the presence of an ELB, I extend the baseline model by allowing for limited attention to inflation and the output gap. Output gap expectations are given by

$$y_{t+1|t}^{gap,e} = y_{t|t-1}^{gap,e} + \gamma^y \left(y_t^{gap} - y_{t|t-1}^{gap,e} \right). \tag{20}$$

As I show in Appendix E.1, limited attention with respect to the output gap exacerbates attention traps. The economy remains stuck at the ELB for longer than in the case of rational expectations and, additionally, inflation, inflation expectations, and now also the output gap stay below their initial values very persistently. The reason for this is that make-up policies, such as forward guidance, not only work through inflation expectations but also through output gap expectations (via the IS equation). When households, however, form their expectations according to (20), output gap expectations become backward looking, which implies that these make-up policies do not stimulate output gap expectations any longer. Thus, the ELB becomes longer lasting.

To prevent these long periods at the ELB, I show in Appendix E.1 that the optimal inflation target becomes even higher than under rational output gap expectations and welfare deteriorates. Higher attention to the output gap serves a similar role as higher attention to inflation: the optimal inflation target is lower and welfare is higher at higher attention levels. These results hold independently of whether attention to inflation is higher than attention to

the output gap (which was likely to be the case before the 1990s) or lower (which was the case before the Covid-19 pandemic).

4.3.3 The Role of the Phillips Curve

Ascari and Haber (2022) show that in high-inflation environments, the price level becomes more responsive to monetary policy shocks. Similarly, Alvarez et al. (2019) find that prices are adjusted more frequently in such high-inflation environments (see also Alexandrov (2022)). Thus, by inducing a higher average inflation rate, i.e., by raising the inflation target, the price-setting behavior of firms may change and thus, affect the optimal average inflation rate. To understand the implications of such changes in price-setting behavior, I now extend the analysis along three dimensions. First, I consider the case of a permanently steeper slope of the Phillips Curve, i.e., a higher κ in equation (7), which may arise due to a lower price adjustment cost. Second, I let the slope of the Phillips Curve to be time-varying, i.e., I model κ as a function of π_t. Third, I solve for the optimal policy for the case in which trend inflation is positive.[20] Table 5 presents the results, which I now discuss in more detail.

Table 5: The Role of the Phillips Curve for Optimal Policy

	Inflation Target		Welfare	
	$\gamma = 0.3$	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.1$
Baseline	1.06%	1.82%	-0.0049	-0.0096
Higher κ				
Fixed χ	0.95%	2.52%	-0.0041	-0.0173
Higher χ	1.02%	2.98%	-0.0045	-0.0239
Time-varying κ_t	1.09%	1.83%	-0.0050	-0.0097
Positive trend inflation	1.04%	2.02%	-0.0029	-0.0069

Notes: This table shows the implications of different Phillips Curve specifications for the optimal inflation target and welfare, for $\gamma = 0.3$ and $\gamma = 0.1$.

For the first case, I set the slope of the Phillips Curve κ to a permanently higher level. This may reflect that in an economy with a higher inflation target, and a higher average inflation rate, the cost of adjusting prices is lower. I set κ to 1.5 times the value I use in my baseline calibration (the baseline calibration sets $\kappa = 0.057$ as in Adam and Billi (2006)). The row in Table 5 labeled *Fixed χ* shows the implications of that change for the optimal inflation target and for welfare for the case in which attention to inflation is relatively high ($\gamma = 0.3$) and in which it is relatively low ($\gamma = 0.1$). The row labeled *Higher χ* shows the results for the case in which the welfare weight on the output gap is adjusted accordingly.[21] When comparing it to the

---

[20]I use trend inflation and steady-state inflation interchangeably.

[21]The weight on the output gap is given by $\chi = \frac{\kappa}{\epsilon}$, where ϵ denotes the price elasticity of demand for the

baseline calibration (row labeled *Baseline*), we see that the optimal inflation target and welfare are not much affected in the case of $\gamma = 0.3$ by these changes. In the case of low attention, $\gamma = 0.1$, however, the optimal inflation target substantially increases by 0.7-1.16 percentage points and welfare decreases. The reason for this increase in the optimal inflation rate is that when the Phillips Curve is steep, fluctuations in the output gap translate into larger changes in inflation. When the economy is pushed to the ELB by adverse demand shocks, output gap decreases and hence, pushes inflation substantially down. Thus, the real rate is relatively high and hence, demand remains low. When attention is low, it takes longer for inflation expectations to recover after such a downturn and the economy remains at or close to the ELB for longer. To prevent this, it is optimal for the central bank to induce a higher inflation rate ex-ante which reduces the likelihood of reaching the ELB in the first place.

In the second case, I allow κ_t to be time-varying and to be a function of inflation π_t. In particular, I assume the following functional form:

$$\kappa_t = \bar{\kappa} + \kappa_1 \pi_t,$$

where I set $\bar{\kappa}$ to my baseline value of 0.057 and set κ_1 to 0.1. This captures the idea that the Phillips Curve steepens when inflation is high in a reduced form way. The row *Time-varying* κ_t shows that this has barely an effect on the optimal inflation rate and welfare. The optimal inflation rate slightly increases, but the quantitative effects are very small.

In the third case, I allow for non-zero trend inflation (see Appendix C for details on the derivations, which follow Ascari and Rossi (2012)). With Rotemberg price-adjustment costs (Rotemberg, 1982), this gives rise to the aggregate IS equation and Phillips Curve:

$$\widehat{y}_t^{gap} = E_t \widehat{y}_{t+1}^{gap} + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2} \left[\pi_t - \pi_{t+1|t}^e \right] - \left(i_t - \pi_{t+1|t}^e - r_t^n \right)$$

$$\pi_t = \zeta \left[\frac{\epsilon(1 + \nu)}{\psi} \widehat{y}_t^{gap} + \beta \bar{\Pi}^2 \pi_{t+1|t}^e + u_t + \Xi \pi_{t+1|t}^e \right]$$

where

$$\zeta \equiv \frac{1}{\bar{\Pi}(2\bar{\Pi} - 1) + \frac{\beta \psi \bar{\Pi}(\bar{\Pi}-1)^2}{1 - \frac{\psi}{2}(\bar{\Pi}-1)^2} + \frac{\epsilon \bar{\Pi}(\bar{\Pi}-1)}{1 - \frac{\psi}{2}(\bar{\Pi}-1)^2}}$$

$$\Xi \equiv \beta \bar{\Pi}(\bar{\Pi} - 1) \left[1 + \frac{\psi(\bar{\Pi} - 1)}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2} \right].$$

Following Ascari and Rossi (2012), I set the demand elasticity $\epsilon = 10$. Given $\kappa = 0.057$, this

---

intermediate goods (Adam and Billi, 2006).

implies a price-adjustment cost parameter $\psi = 350.87$.[22] When trend inflation is zero, $\bar{\Pi} = 1$, it follows that $\zeta = 1$ and $\Xi = 0$, so that we are back to the baseline case analyzed above. Note, that π_t now denotes inflation in deviations from trend inflation (similarly, in the welfare objective (12)). The last row in Table 5 labeled *Positive trend inflation* shows the policy implications of trend inflation of 0.5%. We see that the optimal inflation rate is barely affected for the case of high attention ($\gamma = 0.3$) but increases slightly at low levels of attention ($\gamma = 0.1$). However, the increase in the average inflation rate is less than trend inflation.

Overall these results highlight that the main policy implications of my analysis turn out to be robust to accounting for changing price-setting behavior in high-inflation environments: the optimal inflation target is substantially higher under limited attention to inflation and increases as attention falls.

4.3.4   Time-Varying Attention

Up to now, I assumed that attention does not respond to short-term changes in the economy and therefore, compared economies with different degrees of attention but in which attention was time invariant within economy. To analyze how the results change, when attention responds to short-run changes, I impose that attention takes the form

$$\gamma_t = \bar{\gamma} + \gamma_1 \pi_t.$$

I set $\bar{\gamma}$ to an intermediate value of 0.2 and solve for the optimal policy for the cases $\gamma_1 = 0.3$ and $\gamma_1 = -0.3$. Table 6 shows that when attention increases with inflation ($\gamma_1 = 0.3$), the optimal inflation target increases and welfare decreases. In contrast, when attention decreases with inflation ($\gamma_1 = -0.3$), the optimal inflation rate decreases and welfare increases. When the economy is pushed to the ELB due to an adverse demand shock, inflation decreases. In the case of $\gamma_1 = 0.3$, attention therefore then decreases as well which implies that inflation expectations tend to be lower for longer, keeping actual inflation low, too. Thus, the economy is likely to stay at the ELB for longer. In order to prevent this, it is therefore optimal to induce a higher average inflation rate which makes the ELB less likely to be binding. However, higher average inflation is costly from a welfare perspective, and hence, welfare is lower in that case. When $\gamma_1 = -0.3$, the opposite is the case, and therefore, the optimal inflation rate is lower and welfare is higher.

---

[22]In general, there would be an additional term with the expected change of the output gap in the Phillips Curve, but this term drops out in my case, because I set $\varphi = 1$ (see Appendix C for details).

Table 6: Time-Varying Attention to Inflation

	Inflation Target	Welfare
Baseline	1.20%	-0.0060
$\gamma_1 = 0.3$	1.32%	-0.0069
$\gamma_1 = -0.3$	0.98%	-0.0046

Notes: This table shows the implications of time-varying attention to inflation for the optimal inflation target and welfare.

4.3.5 Full Attention

There are two main assumptions underlying the law of motion of inflation expectations (9). First, the assumption that agents perceive inflation to follow an AR(1) process. Second, that paying attention to inflation is costly and thus, their attention is limited. To disentangle these two effects, Figure 15 in Appendix E.5 shows the implications of shutting down the second channel, i.e., if we set $\gamma = 1$. Put differently, how much of the results in previous sections is solely due to the misperception of the law of motion of inflation? Panel (a) shows that the optimal average inflation rate is practically identical to the one under rational expectations. Thus, as argued above, it is really the lack of attention that pushes up the optimal inflation rate, whereas the assumption of a misperceived law of motion is rather innocuous from this perspective.

Panel (b) in Figure 15 reports the inflation volatility under Ramsey optimal policy for different levels of attention, including the full attention case, $\gamma = 1$. Inflation is substantially more volatile than under rational expectations. Thus, while the misperception of the law of motion of inflation is rather inconsequential for the optimal inflation target, it predicts a higher inflation volatility. Indeed, there is a trade-off. On the one hand, higher attention increases the response of inflation to shocks. On the other hand, if agents are fully attentive, inflation also reacts more strongly to changes in expected future output. Thus, to achieve a certain effect on today's inflation rate, the policymaker is required to make smaller promises about its policies in the future which in turn stabilizes inflation and output already today. While the first channel dominates at lower levels of γ, the second effect pushes inflation volatility down at higher levels of γ.

Finally, panel (c) in Figure 15 shows the welfare implications of the misperceived law of motion. Welfare is slightly more negative when agents have a misperceived law of motion of inflation compared to fully rational agents. Given the results in panels (a) and (b), we see that these additional welfare losses are mainly due to increased inflation volatility rather than its level.

5 Conclusion

With the stabilization of inflation in advanced economies since the Great Inflation period, inflation has become less important in people's lives. In this paper, I quantify this using a limited-attention model of inflation expectations. In line with this model, I show that attention to inflation decreased together with inflation volatility and inflation persistence since the 1970s. Especially in the period between 2010 and 2020, the general public's attention to inflation was close to zero.

For monetary policy the decline in attention was desirable at first, since lower attention stabilizes inflation expectations and hence, stabilizes actual inflation. With the outbreak of the Great Recession and nominal rates at their lower bound, however, managing inflation expectations became a central tool for monetary policy. But managing inflation expectations is difficult when people are inattentive.

The optimal policy response is a substantial increase in the inflation target. This increases the average nominal rate and thus, binding ELB periods become less likely. The cost of this increase in inflation, however, outweighs the stabilization benefits of lower attention. Lower attention, therefore, decreases welfare if we account for the lower bound. This stands in stark contrast to the case without an ELB in which case lower attention leads to welfare gains through the stabilization of inflation expectations and inflation. My paper thus shows that accounting for the ELB is crucial when assessing the role of the public's attention to inflation.

References

ADAM, K. (2007): "Optimal monetary policy with imperfect common knowledge," *Journal of Monetary Economics*, 54, 267–301.

ADAM, K. AND R. M. BILLI (2006): "Optimal monetary policy under commitment with a zero bound on nominal interest rates," *Journal of Money, Credit and Banking*, 1877–1905.

ADAM, K. AND M. PADULA (2011): "Inflation dynamics and subjective expectations in the United States," *Economic Inquiry*, 49, 13–25.

ADAM, K., O. PFÄUTI, AND T. REINELT (2022): "Subjective Housing Price Expectations, Falling Natural Rates and the Optimal Inflation Target," *Working paper*.

AFROUZI, H., S. Y. KWON, A. LANDIER, Y. MA, AND D. THESMAR (2023): "Overreaction in expectations: Evidence and theory," *The Quarterly Journal of Economics*.

AFROUZI, H. AND C. YANG (2021): "Dynamic Rational Inattention and the Phillips Curve," *Working paper*.

ALEXANDROV, A. (2022): "The Effects of Trend Inflation on Aggregate Dynamics and Monetary Stabilization," *Working paper*.

ALVAREZ, F., M. BERAJA, M. GONZALEZ-ROZADA, AND P. A. NEUMEYER (2019): "From hyperinflation to stable prices: Argentina's evidence on menu cost models," *The Quarterly Journal of Economics*, 134, 451–505.

ANDRADE, P., J. GALÍ, H. LE BIHAN, AND J. MATHERON (2019): "The Optimal Inflation Target and the Natural Rate of Interest," *Brookings Papers on Economic Activity*, 2019(2), 173–255.

ANGELETOS, G.-M., Z. HUO, AND K. A. SASTRY (2021): "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 35, 1–86.

ANGELETOS, G.-M. AND C. LIAN (2018): "Forward guidance without common knowledge," *American Economic Review*, 108, 2477–2512.

ASCARI, G. AND T. HABER (2022): "Non-linearities, state-dependent prices and the transmission mechanism of monetary policy," *The Economic Journal*, 132, 37–57.

ASCARI, G. AND L. ROSSI (2012): "Trend inflation and firms price-setting: Rotemberg versus Calvo," *The Economic Journal*, 122, 1115–1141.

BALL, L., N. G. MANKIW, AND R. REIS (2005): "Monetary policy for inattentive economies," *Journal of Monetary Economics*, 52, 703–725.

BENATI, L. (2008): "Investigating inflation persistence across monetary regimes," *The Quarterly Journal of Economics*, 123, 1005–1060.

BENIGNO, P. AND L. PACIELLO (2014): "Monetary policy, doubts and asset prices," *Journal of Monetary Economics*, 64, 85–98.

BHANDARI, A., J. BOROVIČKA, AND P. HO (2022): "Survey data and subjective beliefs in business cycle models," *Working paper.*

BLUNDELL, R. AND S. BOND (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115–143.

BRACHA, A. AND J. TANG (2023): "Inflation levels and (in) attention," *Working Paper.*

BRANDAO-MARQUES, L., M. CASIRAGHI, G. GELOS, G. KAMBER, R. MEEKS, ET AL. (2021): "Negative interest rates: Taking stock of the experience so far," *IMF Departmental Papers*, 3.

CANDIA, B., O. COIBION, AND Y. GORODNICHENKO (2021): "The Inflation Expectations of US Firms: Evidence from a New Survey," *Working paper.*

CANOVA, F. (2007): "G-7 inflation forecasts: Random walk, Phillips curve or what else?" *Macroeconomic Dynamics*, 11, 1.

CARROLL, C. D. (2003): "Macroeconomic expectations of households and professional forecasters," *the Quarterly Journal of economics*, 118, 269–298.

CAVALLO, A., G. CRUCES, AND R. PEREZ-TRUGLIA (2017): "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments," *American Economic Journal: Macroeconomics*, 9, 1–35.

CLARIDA, R., J. GALÍ, AND M. GERTLER (1999): "The Science of Monetary Policy: Evidence and Some Theory," *Journal of Economic Literature*, 37, 1661–1707.

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

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

——— (2015b): "Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation," *American Economic Journal: Macroeconomics*, 7, 197–232.

COIBION, O., Y. GORODNICHENKO, E. S. KNOTEK, AND R. SCHOENLE (2023): "Average inflation targeting and household expectations," *Journal of Political Economy Macroeconomics*, 1, 403–446.

COIBION, O., Y. GORODNICHENKO, AND M. WEBER (2022): "Monetary policy communications and their effects on household inflation expectations," *Journal of Political Economy*, 130, 1537–1584.

COIBION, O., Y. GORODNICHENKO, AND J. WIELAND (2012): "The optimal inflation rate in New Keynesian models: should central banks raise their inflation targets in light of the zero lower bound?" *Review of Economic Studies*, 79, 1371–1406.

CONSTANCIO, V. (2015): "Understanding inflation dynamics and monetary policy," in *Speech at the Jackson Hole Economic Policy Symposium*, vol. 29.

DEL NEGRO, M., M. LENZA, G. E. PRIMICERI, AND A. TAMBALOTTI (2020): "What's up with the Phillips Curve?" *Working paper.*

D'ACUNTO, F., D. HOANG, AND M. WEBER (2022): "Managing households' expectations with unconventional policies," *The Review of Financial Studies*, 35, 1597–1642.

EGGERTSSON, G. AND M. WOODFORD (2003): "The Zero Interest-Rate Bound and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, (1), 139–211.

FAUST, J. AND J. H. WRIGHT (2013): "Forecasting inflation," in *Handbook of economic forecasting*, Elsevier, vol. 2, 2–56.

FULTON, C. AND K. HUBRICH (2021): "Forecasting US inflation in real time," *Econometrics*, 9, 36.

GABAIX, X. (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

——— (2020): "A behavioral New Keynesian model," *American Economic Review*, 110, 2271–2327.

GALÍ, J. (2015): *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press.

GÁTI, L. (2022): "Monetary Policy & Anchored Expectations An Endogenous Gain Learning Model," *Working paper.*

HAMILTON, J. D. (1994): *Time series analysis*, Princeton university press.

JØRGENSEN, P. L. AND K. J. LANSING (2023): "Anchored Inflation Expectations and the Slope of the Phillips Curve," .

JURADO, K. (2023): "Rational inattention in the frequency domain," *Journal of Economic Theory*, 208, 105604.

KORENOK, O., D. MUNRO, AND J. CHEN (2022): "Inflation and attention thresholds," *Working paper.*

LAMLA, M. J. AND S. M. LEIN (2014): "The role of media for consumers' inflation expectation formation," *Journal of Economic Behavior & Organization*, 106, 62–77.

LINK, S., A. PEICHL, C. ROTH, AND J. WOHLFART (2023): "Attention to the Macroeconomy," *Working Paper.*

MAĆKOWIAK, B., F. MATĚJKA, AND M. WIEDERHOLT (2018): "Dynamic rational inattention: Analytical results," *Journal of Economic Theory*, 176, 650–692.

——— (2023): "Rational inattention: A review," *Journal of Economic Literature*, 61, 226–273.

MACKOWIAK, B. AND M. WIEDERHOLT (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99, 769–803.

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

MARCET, A. AND R. MARIMON (2019): "Recursive contracts," *Econometrica*, 87, 1589–1631.

MATĚJKA, F. AND A. MCKAY (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 105, 272–98.

NEWEY, W. K. AND K. D. WEST (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation," *Econometrica*, 55, 703–708.

NICKELL, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica: Journal of the econometric society*, 1417–1426.

PACIELLO, L. AND M. WIEDERHOLT (2014): "Exogenous information, endogenous information, and optimal monetary policy," *Review of Economic Studies*, 81, 356–388.

PFAJFAR, D. AND E. SANTORO (2013): "News on inflation and the epidemiology of inflation expectations," *Journal of Money, Credit and Banking*, 45, 1045–1067.

PFÄUTI, O. (2023): "The Inflation Attention Threshold and Inflation Surges," *arXiv preprint arXiv:2308.09480*.

PFÄUTI, O. AND F. SEYRICH (2022): "A behavioral heterogeneous agent New Keynesian model," *Working Paper*.

ROTEMBERG, J. J. (1982): "Sticky Prices in the United States," *Journal of Political Economy*, 90, 1187–1211.

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

VELLEKOOP, N. AND M. WIEDERHOLT (2019): "Inflation expectations and choices of households," *Working paper*.

WEBER, M., S. SHEFLIN, T. ROPELE, R. LLUBERAS, S. FRACHE, B. MEYER, S. KUMAR, Y. GORODNICHENKO, D. GEORGARAKOS, O. COIBION, ET AL. (2023): "Tell Me Something I Don't Already Know: Learning in Low and High-Inflation Settings," *Working Paper*.

WIEDERHOLT, M. (2015): "Empirical properties of inflation expectations and the zero lower bound," *Working paper*.

WOODFORD, M. (2003): "Interest and prices," *Princeton University Press, Princeton, NJ*.

Online Appendix

A   A Limited-Attention Model of Inflation Expectations

In this section, I derive the expectations-formation process under limited attention sketched in Section 2. The agent believes that (demeaned) inflation tomorrow, π', depends on (demeaned) inflation today, π, as follows

$$\pi' = \rho_\pi \pi + \nu,$$

where $\rho_\pi \in [0,1]$ denotes the perceived persistence of inflation and $\nu \sim i.i.N.(0, \sigma_\nu^2)$. Inflation in the current period is unobservable, so before forming an expectation about future inflation, the agent needs to form an expectation about today's inflation. I denote this nowcast $\widetilde{\pi}$, and the resulting forecast about next period's inflation $\pi^e = \rho_\pi \widetilde{\pi}$. Given her beliefs, the full-information forecast π^{e*} is

$$\pi^{e*} \equiv \rho_\pi \pi.$$

But since π is not perfectly observable, the actual forecast will deviate from the full-information forecast. Deviating, however, is costly, as this causes the agent to make mistakes in her decisions.

The agent's choice is not only about how to form her expectations given certain information, but about how to choose this information optimally, while taking into account how this will later affect her forecast. That is, she chooses the form of the signal s she receives about current inflation. Since acquiring information is costly, it cannot be optimal to acquire different signals that lead to an identical forecast. Due to this one-to-one relation of signal and forecast, we can directly work with the joint distribution of π^e and π, $f(\pi^e, \pi)$, instead of working with the signal.

Let $U(\pi^e, \pi)$ denote the negative of the loss that is incurred when the agent's forecast deviates from the forecast under full information, and $C(f)$ the cost of information. Then, the agent's problem is given by

$$\max_f \int U(\pi^e, \pi) f(\pi^e, \pi) d\pi d\pi^e - C(f) \tag{21}$$

$$\text{subject to } \int f(\pi^e, \pi) d\pi^e = g(\pi), \text{ for all } \pi,$$

where $g(\pi)$ is the agent's prior, which is assumed to be Gaussian; $\pi \sim N\left(\hat{\pi}, \sigma_\pi^2\right)$. $C(.)$ is the cost function that captures how costly information acquisition is. It is linear in *mutual information* $I(\pi; \pi^e)$, i.e., the expected reduction in entropy of π due to knowledge of π^e:

$$C(f) = \lambda I(\pi; \pi^e) = \lambda \left(H(\pi) - E\left[H(\pi|\pi^e)\right] \right),$$

where $H(x) = -\int f(x)log(f(x))dx$ is the entropy of x and λ is a parameter that measures the cost of information.

The objective function $U(.)$ is assumed to be quadratic:

$$U(\pi^e, \pi) = -r\left(\rho_\pi \pi - \pi^e\right)^2,$$

where r measures the stakes of making a mistake.[23,24]

In this setup, Gaussian signals are optimal (and in fact the unique solution, see Matějka and McKay (2015)). The optimal signal thus has the form

$$s = \pi + \varepsilon,$$

with $\varepsilon \sim i.i.N.(0, \sigma_\varepsilon^2)$.[25] The problem (21) now reads

$$\max_{\sigma_{\pi|s}^2 \leq \sigma_\pi^2} E_\pi \left[E_s \left[-r\rho_\pi^2 \left(\pi - E[\pi|s]\right)^2 \right] \right] - \lambda I(\pi; \pi^e) = \max_{\sigma_{\pi|s}^2 \leq \sigma_\pi^2} \left(-r\rho_\pi^2 \sigma_{\pi|s}^2 - \frac{\lambda}{2} log \frac{\sigma_\pi^2}{\sigma_{\pi|s}^2} \right). \qquad (22)$$

The optimal forecast is given by $\pi^e = \rho_\pi E\left[\pi|s\right]$, and Bayesian updating implies

$$\pi^e = \rho_\pi \left(1 - \gamma\right) \hat\pi + \rho_\pi \gamma s, \qquad (23)$$

where $\gamma = 1 - \frac{\sigma_{\pi|s}^2}{\sigma_\pi^2} \in [0, 1]$ measures how much attention the agent pays to inflation, and $\hat\pi$ denotes the prior mean of π.

An equivalent way of writing γ is

$$\gamma = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_\varepsilon^2}. \qquad (24)$$

Now, since the agent *chooses* the level of attention, we can re-formulate (22) as

$$\max_{\gamma \in [0,1]} \left(-r\rho_\pi^2(1 - \gamma)\sigma_\pi^2 - \frac{\lambda}{2} log \frac{1}{1 - \gamma} \right). \qquad (25)$$

---

[23]A quadratic loss function is usually derived from a second-order approximation of the household's utility function or the firm's profit function (see, e.g., Mackowiak and Wiederholt (2009)).

[24]These stakes (or also the information cost parameter λ) can be interpreted as a way to incorporate other variables to which the agent might pay attention. For example, a household might not only want to forecast inflation but also her own income stream going forward. In this case, a smaller r could capture an increase in her idiosyncratic income volatility. Thus, paying attention to inflation is relatively less beneficial, as the relative importance of her idiosyncratic income increases. Such an interpretation also explains why professional forecasters might not be fully informed about inflation, given that they usually forecast a whole array of variables.

[25]In this case, the entropy becomes $H(x) = \frac{1}{2}log(2\pi e\sigma_x^2)$, where σ_x^2 is the variance of x. Note, that here π denotes the number "pi" and not inflation.

Writing the cost of information relative to the stakes, $\tilde{\lambda} \equiv \frac{\lambda}{r}$, and solving the optimization problem (25) yields the *optimal* level of attention, presented in Lemma 1.

B   Appendix to Empirical Results

Figure 6 shows the main time series that are used in the empirical analyses of Section 2. Apart from the apparent decrease in the level and volatility of inflation as well as inflation expectations, we see that expectations became more and more detached from actual inflation. First, consumer expectations seem to be biased on average in the most recent decades, as can be seen in the lower panel. While these expectations closely tracked inflation in the 70s and 80s, this is not the case anymore.[26] Second, professional forecasters' expectations seem to perform quite well on average. In the last twenty years, however, they barely react to actual changes in inflation anymore. Overall, these observations suggest that attention decreased in the last decades.

Figure 6: Inflation and Inflation Expectations
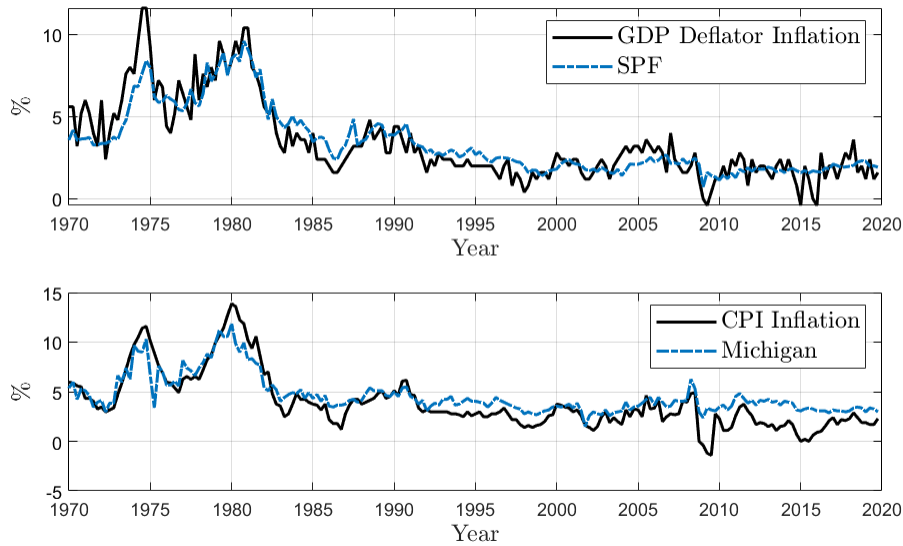
Note: This figure shows the raw time series of inflation, as well as survey expectations about future inflation. Everything is in annualized percentages.

Table 7 shows the summary statistics, for the period before and after the 1990s, separately. For professional forecasters, the perceived persistence is higher than the actual one. This is especially the case when the actual persistence is relatively low, as was the case after 1990. Afrouzi et al. (2023) document a similar finding in an experimental setting. This might point towards lower attention since the 1990s. Note, that in the empirical analysis I account for changes in the perceived persistence.

---

[26]In the empirical analysis I account for this mean bias by including an intercept in the regressions.

Table 7: Summary Statistics

	GDP Deflator Inflation		SPF Expectations	
	1968-1990	1990-2020	1968-1990	1990-2020
Mean (%)	5.44	2.00	5.18	2.15
Std. Dev. (%)	2.43	0.90	1.87	0.63
Persistence	0.84	0.55	0.93	0.92
	CPI Inflation		Consumer Expectations	
	1968-1990	1990-2020	1968-1990	1990-2020
Mean (%)	6.09	2.42	6.00	3.62
Std. Dev. (%)	3.00	1.26	2.17	0.68
Persistence	0.96	0.77	0.85	0.70

Note: This table shows the summary statistics of the data. The upper panel shows the statistics for the quarter-on-quarter GDP deflator inflation (left) and the corresponding inflation expectations from the Survey of Professional Forecasters (right). The lower panel shows the year-on-year CPI inflation (left) and the corresponding inflation expectations from the Survey of Consumers from the University of Michigan. All data are annualized.

B.1 Robustness and Additional Evidence

In this section, I show that the empirical results are robust along several dimensions.

Additional Data Sources

In Table 8, I show how attention changed over time for different data sources. The first two columns show the results for the Greenbook forecasts, columns 3-4 for the Livingston Survey, and columns 5-6 and 7-8 are for CPI forecasts from the SPF instead of forecasts about the GDP deflator. As in the main text, I use two different estimators. First, the Blundell-Bond estimator (columns 5-6) and pooled OLS (columns 7-8). All standard errors are robust with respect to heteroskedasticity and serial correlation. We see that the main finding of lower attention in inflation expectations in the period after 1990 compared to the period before is robust to these changes in the data source and/or exact variable.

Table 8: Regression Results of Equation (4)

	Greenbook		Livingston		SPF CPI BB		SPF CPI OLS	
	< 1990	≥ 1990	< 1990	≥ 1990	< 1990	≥ 1990	< 1990	≥ 1990
$\widehat{\gamma}$	0.39	0.24	0.28	0.17	0.36	0.23	0.17	0.13
s.e.	(0.0851)	(0.0715)	(0.0554)	(0.0624)	(0.1444)	(0.0328)	(0.0409)	(0.0142)
N	84	100	83	61	550	3,577	550	3,577

Note: This table shows the results from regression (4) for different data sources. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Different Sample Splits

Table 1 in the main body of the paper shows that attention to inflation declined by focusing on a sample split in 1990. To show that this is robust to the exact split point, Tables 9 and 10 show that the result holds when splitting the sample in 1985 or 1995, respectively. In fact, the decline in attention is even somewhat more pronounced when splitting the sample in 1985. This is in line with the theoretical prediction of the limited-attention model. Namely, the period between 1985 and 1990 was a period of relatively low and stable inflation compared to the period pre 1985 (see Figure 6), and thus, a period in which the model would predict a relatively low level of attention.

Table 9: Regression Results of Equation (4), pre 1985 vs. post 1985

	Professional Forecasters				Consumers			
	Blundell Bond		Pooled OLS		Averages		Median	
	< 1985	≥ 1985	< 1985	≥ 1985	< 1985	≥ 1985	< 1985	≥ 1985
$\widehat{\gamma}$	0.75	0.37	0.45	0.25	0.77	0.31	0.50	0.26
s.e.	(0.1247)	(0.0399)	(0.0403)	(0.0338)	(0.1688)	(0.0811)	(0.0955)	(0.0561)
N	1914	3887	1914	3887	64	140	27	140

Note: This table shows the results from regression (4) for professional forecasters (SPF) as well as for consumers. For the SPF, I use the Blundell and Bond (1998) (BB) estimator (first two columns), as well as pooled OLS (columns 3-4). For the Survey of Consumer, I consider average expectations (columns 5-6) and median expectations (columns 7-8). The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Table 10: Regression Results of Equation (4), pre 1995 vs. post 1995

	Survey of Professional Forecasters				Survey of Consumers			
	Blundell Bond		Pooled OLS		Averages		Median	
	< 1995	≥ 1995	< 1995	≥ 1995	< 1995	≥ 1995	< 1995	≥ 1995
$\widehat{\gamma}$	0.70	0.41	0.44	0.21	0.72	0.27	0.43	0.22
s.e.	(0.0907)	(0.0654)	(0.0379)	(0.0344)	(0.1473)	(0.0962)	(0.0819)	(0.0654)
N	2708	3093	2708	3093	104	100	67	100

Note: This table shows the results from regression (4) for professional forecasters (SPF) as well as for consumers. For the SPF, I use the Blundell and Bond (1998) (BB) estimator (first two columns), as well as pooled OLS (columns 3-4). For the Survey of Consumer, I consider average expectations (columns 5-6) and median expectations (columns 7-8). The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Different Specifications of the BB Estimator

In the baseline estimation, reported in Table 1, I included all potential lags for the Blundell-Bond estimation. To show that the results are robust to this specification, I show in Table 11 that for maximum lag lengths of 20 and 10 periods, the estimated attention parameter $\widehat{\gamma}$ is in all cases higher before 1990 compared to the period after 1990.

Table 11: Different Maximum Lag Lengths

	All Lags		20 Lags		10 Lags	
	< 1990	≥ 1990	< 1990	≥ 1990	< 1990	≥ 1990
$\widehat{\gamma}$	0.70	0.41	0.74	0.51	0.84	0.69
s.e.	(0.1005)	(0.0522)	(0.1086)	(0.0632)	(0.1247)	(0.1127)
N	2235	3566	2235	3566	2235	3566

Note: This table shows the results from regression (4) for different numbers of lags included in the BB estimation. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Time Fixed Effects

To account for potential changes in trend inflation, I include time-fixed effects in regression (4). To do so, recall that (4) is given by

$$\pi^e_{t+1|t,i} = \beta_i + \beta_1 \pi^e_{t|t-1,i} + \beta_2 \left(\pi_t - \pi^e_{t|t-1,i} \right) + u_{i,t}. \tag{26}$$

To include time fixed effects, I first compute a period-specific persistence parameter, ρ_π. Note, that in (26), β_1 measures this persistence. Therefore, I subtract $\widehat{\rho}_\pi \pi^e_{t|t-1,i}$ from both sides and then to directly estimate γ, I further divide both sides by $\widehat{\rho}_\pi$:

$$\frac{\pi^e_{t+1|t,i} - \widehat{\rho}_\pi \pi^e_{t|t-1,i}}{\widehat{\rho}_\pi} = \delta_i + d_t + \gamma \left(\pi_t - \pi^e_{t|t-1,i} \right) + v_{i,t}, \tag{27}$$

where d_t captures time-fixed effects, $\delta_i = \frac{\beta_i}{\widehat{\rho}_\pi}$ and $v_{i,t} = \frac{u_{i,t}}{\widehat{\rho}_\pi}$. I do this transformation for the period before and after 1990 separately. Note, that this transformation also deals with the endogeneity problem explained in Section 2.

The estimated attention levels are 0.75 (s.e. 0.0327) for the period before 1990 and 0.61 (s.e. 0.0295) after 1990 if I use the first-order autocorrelation of expected inflation as my measure of ρ_π. If I use the estimate of β_1 from equation (27) as my measure of ρ_π, the estimated attention before the 1990s is 0.68 (s.e. 0.0252) and the one after the 1990s is 0.46 (s.e. 0.0242). Thus, we see that the decrease in attention is robust to controlling for time-fixed effects, even though the decline is somewhat muted.

When using the first-order autocorrelation of expected inflation as my measure of ρ_π, estimating equation (5) in this way, delivers a point estimate of 0.06 (s.e. 0.0111) that is statistically significant on all conventional significance levels. The estimate for ζ in regression (6) is 0.23 (s.e. 0.0220), statistically significant on all conventional significance levels. When using $\hat{\beta}_1$ from (27) as the measure of ρ_π, the point estimate of β in equation (5) is 0.06 (s.e. 0.0074) and the estimate of ζ in (6) is 0.29 (s.e. 0.0306), both statistically significant on all conventional levels of significance. Thus, the positive relationships between attention and volatility, as well as between attention and inflation persistence, are robust to controlling for time fixed effects.

Professional Forecasters in the Aggregate

When estimating attention of professional forecastors' average expectations instead of individual ones, we obtain a value of 0.24 (s.e. 0.0481) for the period before 1990 and of 0.09 (s.e. 0.0353) after 1990. Consistent with the main results, attention substantially decreased in recent decades and is about half after 1990 compared to before.

Estimating regression (5) on aggregate SPF data delivers a coefficient of 0.15 (p-value of 0.000) and the estimate of ζ in regression (6) is 0.69 (p-value of 0.000). Thus, the results reported in the main text are robust.

Joint Regressions

Instead of running regressions (5) and (6) separately, I estimate

$$\widehat{\gamma}_t = \alpha + \beta\widehat{\sigma}_{\pi,t} + \zeta\widehat{\rho}_{\pi,t} + u_t. \tag{28}$$

Table 12 shows that the results are robust to this change in specification.

Table 12: Attention, Inflation Volatility and Inflation Persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.04***	0.05***	0.06***
s.e.	(0.0153)	(0.0128)	(0.0150)
$\widehat{\zeta}$	0.59***	0.65***	0.31***
s.e.	(0.0597)	(0.0499)	(0.0772)
N	165	165	163

Note: This table shows the results of regression (28). Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Controlling for Average Inflation

A potential confounder in regression (28) above is the average level of inflation. Thus, I now control for the average level of inflation, computed as the average inflation rate in the respective 10-year window. In particular, I run the following regression

$$\widehat{\gamma}_t = \alpha + \beta\widehat{\sigma}_{\pi,t} + \zeta\widehat{\rho}_{\pi,t} + \omega\widehat{\overline{\pi}}_t + u_t, \tag{29}$$

where $\widehat{\overline{\pi}}_t$ is the estimated average inflation rate. Table 13 reports the results for the professional forecasters. We see that the volatility and the persistence of inflation are positively related with attention and that these relationships are statistically significant even when controlling for the average level of inflation. The average level of inflation, on the other hand, does not have a positive, statistically-significant, effect on the estimated attention when we control for the volatility and persistence of inflation. These results are consistent with the underlying theoretical model.

Table 13: Controlling for Average Inflation

	Survey of Professional Forecasters	
Estimator	Blundell-Bond	Pooled OLS
$\widehat{\beta}$	0.13***	0.06***
s.e.	(0.0264)	(0.0200)
$\widehat{\zeta}$	0.69***	0.82***
s.e.	(0.1153)	(0.0824)
$\widehat{\omega}$	−0.02**	0.01
s.e.	(0.0095)	(0.0082)
N	165	165

Note: This table shows the results of regression (29). Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Quasi-Panel of Consumers

The Survey of Consumers does not follow consumers over time. Therefore, I could not allow for individual-specific fixed effects but rather consider average and/or median inflation expectations. I now group the survey respondents into four groups, based on their income. The SoC provides data on this starting in the last quarter of 1979.

Table 14 shows the results. The first two columns report the results for the split point in 1990, and the third and fourth column for the split point in 1995. We see that the estimated attention levels using this quasi panel are similar to the ones obtained using average expectations (Table 1).

Table 14: Regression Results of Equation (4), Quasi-Panel

	Survey of Consumers			
	< 1990	≥ 1990	< 1995	≥ 1995
$\widehat{\gamma}$	0.77	0.33	0.70	0.29
s.e.	(0.0933)	(0.0263)	(0.1078)	(0.0289)
N	160	480	240	400

Note: This table shows the results from regression (4), estimated using the Blundell and Bond (1998) estimator, for consumers grouped into four groups, based on their income. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Table 15 shows the results of regressions (5) and (6) (first column), as well as of the joint regression (28), using this quasi panel of consumers. We see that the results are robust and that there is indeed a significantly positive relation between attention and inflation volatility, as well as between attention and inflation persistence.

Table 15: Attention, Inflation Volatility and Inflation Persistence

	Survey of Consumers	
Estimator	Separate	Joint
$\widehat{\beta}$	0.13***	0.13***
s.e.	(0.0106)	(0.0126)
$\widehat{\zeta}$	0.20***	0.12***
s.e.	(0.0787)	(0.0620)
N	121	121

Note: This table shows the results of regressions (5), (6) (first column) and (28) (second column) using a quasi panel of consumers. The attention parameters have been estimated using the BB-estimator. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Volatility and Persistence of Inflation Expectations

Table 16 shows the results of regressions (5) and (6) using the volatility and persistence of inflation expectations instead of actual inflation as independent variables. Standard errors are robust with respect to heteroskedasticity.

Table 16: Attention, Inflation Volatility and Inflation Persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.14***	0.16***	0.13***
s.e.	(0.0153)	(0.0098)	(0.0172)
$\widehat{\zeta}$	1.06***	1.21***	0.23***
s.e.	(0.1272)	(0.0838)	(0.0685)
N	165	165	163

Note: This table shows the results of regressions (5) and (6) using the volatility and persistence of inflation expectations instead of actual inflation as dependent variables. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Window Length

As predicted by the underlying model of optimal information acquisition, I showed that there is indeed a positive relationship between attention to inflation and inflation volatility, as well as between attention and inflation persistence. In the baseline specification, I relied on a rolling-window approach in which every window was 10 years. Tables 17 and 18 show that these results are robust to using different window lengths, namely 5 and 15 years.

Table 17: Attention, Inflation Volatility and Inflation Persistence

|	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	-0.01	0.06***	0.13***
s.e.	(0.1643)	(0.0185)	(0.0411)
$\widehat{\zeta}$	0.73	0.44***	0.40***
s.e.	(0.6731)	(0.0551)	(0.1547)
N	185	185	183

Note: This table shows the results of regression (28) using windows of 5 years each. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Table 18: Attention, Inflation Volatility and Inflation Persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.01	0.01	0.07***
s.e.	(0.0124)	(0.0115)	(0.0136)
$\widehat{\zeta}$	0.90***	1.00***	0.43***
s.e.	(0.0603)	(0.0552)	(0.0706)
N	145	145	143

Note: This table shows the results of regression (28) using windows of 15 years each. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Attention over Time

Figure 7 shows the estimated attention levels, γ, (black-solid line) from the SPF consensus forecasts, together with the volatility of GDP deflator inflation (blue-dashed lines). We clearly see the aforementioned decrease in attention over time, as well as the positive correlation of attention and inflation volatility.

Figure 7: Attention and Inflation Volatility over Time
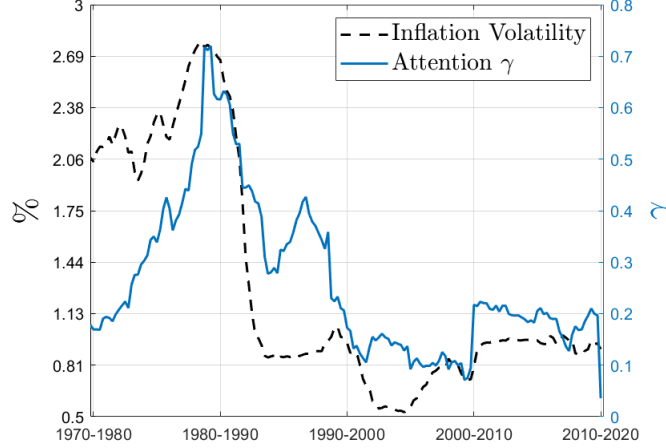
Notes: This figure shows the estimated attention levels, γ, (black-solid line) from the SPF consensus forecasts, together with the volatility of GDP deflator inflation (blue-dashed lines).

AR(2) Beliefs

In the main part of the paper, I assume that agents believe that inflation follows an AR(1). I now show that the main results are unchanged when instead assuming that agents believe that inflation follows an AR(2).

Assume agents have a law of motion of demeaned inflation given by

$$\pi_t = \phi_1 \pi_{t-1} + \phi_2 \pi_{t-2} + \nu_t,$$

and that they receive signals of the form $s_t = \pi_t + \epsilon_t$, and the disturbances ν and ϵ are i.i.d. zero-mean normally-distributed random variables with time-invariant volatilities.[27] Following the arguments in Hamilton (1994), the steady state Kalman filter then yields:

$$\begin{pmatrix} \pi_{t+1|t}^e \\ \pi_{t|t}^e \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \pi_{t|t-1}^e \\ \pi_{t-1|t-1}^e \end{pmatrix} + \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} \left(\pi_t + \epsilon_t - \pi_{t|t-1}^e \right).$$

---

[27]Assuming a signal of the form $s_t = \varphi_1 \pi_t + \varphi_2 \pi_{t-1} + \epsilon_t$ (which would be the optimal signal in the AR(2) case, see Maćkowiak et al. (2018) and Jurado (2023)) would lead to an identification issue because the weights in the signal φ_1 and φ_2 could not be disentangled from the updating weights, k_1 and k_2.

The second equation, shifted one period backwards, yields

$$\pi^e_{t-1|t-1} = \pi^e_{t-1|t-2} + k_2 \left(\pi_{t-1} + \epsilon_{t-1} - \pi^e_{t-1|t-2} \right),$$

which we can then plug into the first equation to obtain an expression for the one-period ahead expectations $\pi^e_{t+1|t}$:

$$\pi^e_{t+1|t} = \phi_1 \pi^e_{t|t-1} + \phi_2 \pi^e_{t-1|t-2} + k_1 \left(\pi_t - \pi^e_{t|t-1} \right) + \underbrace{K_2}_{=\phi_2 k_2} \left(\pi_{t-1} - \pi^e_{t-1|t-2} \right) + u_t, \qquad (30)$$

where k_1 and k_2 are the coefficients in the Kalman gain matrix (denoted by K in Hamilton (1994)), and $u_t = k_1 \epsilon_t + K_2 \epsilon_{t-1}$. I consider household average and median expectations when estimating regression (30), and I include an intercept. To account for serial correlation in the error term, I apply the Newey-West estimator using four lags (Newey and West (1987)). Table 19 shows the results.

Table 19: AR(2) Perceived Law of Motion

| | Michigan Survey | | | |
| | Average Expectations | | Median Expectations | |
	pre 1990	post 1990	pre 1990	post 1990
$\widehat{k}_1$	0.86	0.30	0.446	0.21
s.e.	(0.167)	(0.084)	(0.107)	(0.0599)
$\widehat{K}_2$	-0.35	-0.14	-0.105	-0.0889
s.e.	(0.156)	(0.082)	(0.112)	(0.0529)
$\widehat{\phi}_1$	1.11	0.88	1.09	0.82
s.e.	(0.178)	(0.100)	(0.226)	(0.127)
$\widehat{\phi}_2$	-0.413	-0.136	-0.276	-0.166
s.e.	(0.190)	(0.083)	(0.2203)	(0.085)

Note: This table shows the results of regression (30) for household average and median expectations. Standard errors are robust with respect to heteroskedasticity and serial correlation (Newey-West with 4 lags).

When using the sum of the two updating gains, $k_1 + K_2$, as the measure of attention, we see that attention clearly decreased from the period before the 1990s to the period after the 1990s. Before 1990, the sum of the two updating gains when focusing on average expectations is 0.51 (with s.e. of 0.13, so statistically significantly different from 0 at the 1% significance level). After 1990, this measure of attention decreased to 0.16 (again, statistically significantly different from 0 at the 1% significance level). When focusing on median expectations, the sum of the two before the 1990s is 0.341 and it decreased in the period after 1990 to 0.12. This decrease in attention—measured as how strongly households update their expectations—is consistent with the findings in the main text in Section 2.

Furthermore, the estimates in Table 19 show that none of the estimated coefficients that are arise due to the AR(2) assumption—the estimates for K_2 and ϕ_2—are statistically significant at the 1% level (often, not even at the 5% or 10% level). When estimating attention in exactly the same way as in Section 2, i.e., computing attention as $\frac{k_1}{\phi_1}$, I obtain estimates that are very close to the ones in the main text where I ignore the effects arising from the second lag of inflation in the perceived law of motion. For average expectations, I estimate for the period before 1990 an attention parameter of 0.78 (it was 0.75 in the main text), and for the period after 1990 a value of 0.34 (0.31 in the main text). For median expectations, I obtain attention estimates of 0.41 for the period before 1990 (it is 0.43 in the main text) and 0.255 for the period after 1990 (0.24 in the main text). These findings give empirical support to the assumption that the perceived law of motion for inflation follows an AR(1).

B.2 Other Measures of Attention

In this section, I provide complementary evidence to the one presented in Section 2 based on news coverage, based on the share of survey respondents that answer "I don't know" when asked about their inflation expectations, and based on assessing the accuracy of nowcasts of inflation.

News Coverage of Inflation

Figure 8 shows the relative frequency of the word "inflation" among all words in two major U.S. newspapers (blue-dashed lines), the New York Times (left panel) and the Washington Post (right panel), together with the annual U.S. CPI inflation (black-solid lines). It is evident that news coverage is higher in times of high and volatile inflation as was the case during the 1970s and early 1980s. Moreover, the figure suggests that the public's attention to inflation—proxied here by news coverage—has not always been as low as in recent years, but declined over time.

In Figure 9, we see that a similar picture emerges when looking at the coverage of "inflation" in books, according to *Google Books Ngram Viewer*. In the left panel, we see that "inflation" is covered more frequently in English books written in times of high inflation. But this is not simply a U.S. phenomenon. To see this, I show the same statistic for books written in Spanish for the word "inflación". To contrast this with inflation, the black solid line shows the average inflation (in logs) of the four largest Spanish-speaking countries, weighted by their 2020 population size. These are Argentina, Colombia, Mexico and Spain. Again, we observe that attention to inflation—measured by book coverage—is higher in times of high and volatile inflation.

Figure 8: News Coverage of Inflation

(a) New York Times
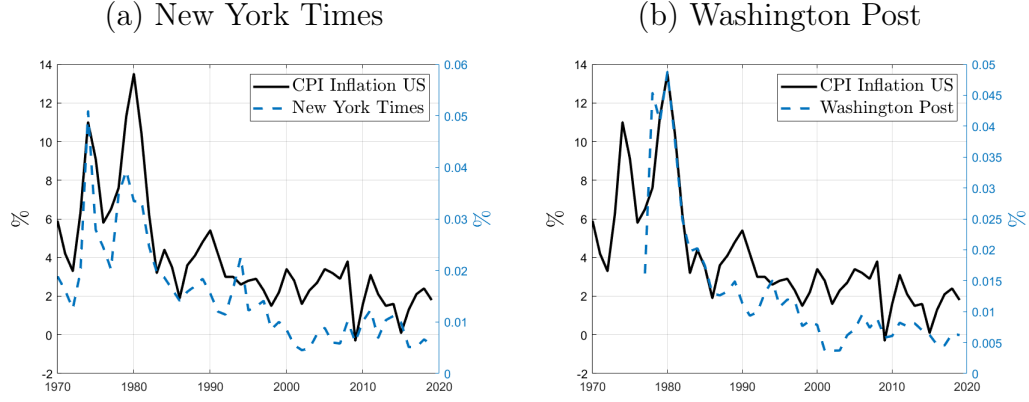
(b) Washington Post

Notes: This figure shows the relative frequency (blue dashed lines, right axis) of the word "inflation" in the New York Times (left) and the Washington Post (right). The black solid line shows annual U.S. CPI inflation (left axis).

Figure 9: Book Coverage of Inflation

(a) Google Books English

(b) Google Books Spanish
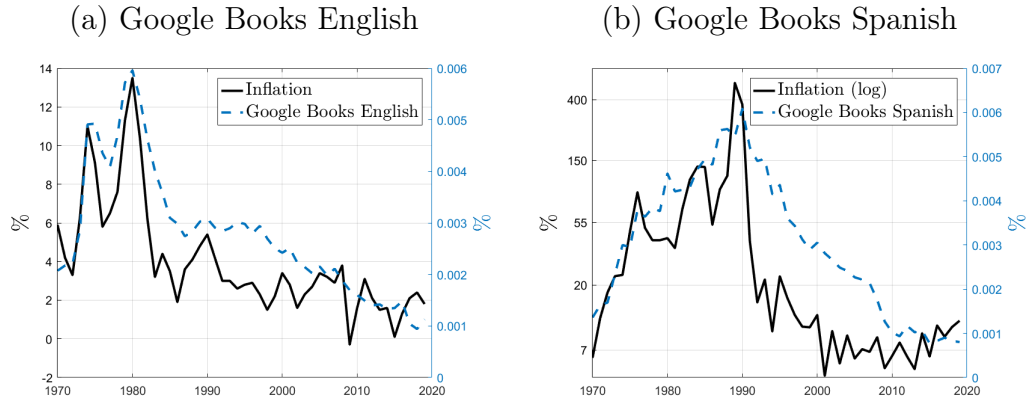
Notes: The blue dashed lines show the frequency of the words "inflation" and "inflación", respectively, in English and Spanish books, according to Google Books Ngram Viewer. The black solid line shows the corresponding inflation rates.

Households answering "I don't know"

Another potential measure of people's inattention is to see how the share of survey respondents that answer the question about their inflation expectations with "I don't know" changes over time. The Michigan survey provides these shares. Following a rolling-windows approach, I compute for each 10-year window the estimated attention parameter $\widehat{\gamma}_t$ as well as the average share of households within these 10 years that say "I don't know". When using average expectations to estimate γ, I find that the two are strongly negatively correlated with a correlation coefficient of -0.4. When regressing the share of "don't know" respondents on the estimated attention parameter, I obtain a regression coefficient of -7.23 (p-value of 0.010). When controlling for the window-specific inflation volatility, autocorrelation and average inflation rate, the regression coefficient is -5.64 with a p-value of 0.016. When using median expectations to compute γ, the results are even slightly stronger. The raw correlation is -0.47, the regression coefficient without controls is -4.29 with a p-value of 0.000, and when adding controls it equals -3.16 with a p-value of 0.013. These results indicate that inattention (as measured by the share of respondents answering "I don't know") is lower in times my measure of attention, γ, is higher. Thus, these findings support the view that my measure of attention indeed captures people's attention to inflation.

Accuracy of Nowcasts

In a setup in which the agent cannot distinguish between a trend and a cyclical component of inflation with time-varying volatilities of these two components, if the trend component's contribution to overall inflation increases, the agent's forecast would become more responsive to current inflation, too, similarly to an increase in the attention parameter γ. To differentiate these two models, I therefore now also consider *nowcasts* of inflation and their accuracy. The optimal attention choice problem presented in Section 2 says that more attentive agents receive more precise signals about current inflation and should therefore make smaller nowcast errors in times of high attention. This prediction is exclusive to the proposed model of attention and does not apply to the alternative model of the trend and cycle component of inflation. Using the nowcasts from the Survey of Professional Forecasters, I now test this prediction of the model. To do this, I take the absolute value of forecast errors of current inflation or the squared forecast errors as my two measures of the accuracy of the forecasters' nowcasts. I then compute the average across all forecasters and estimate a time series of these average forecast errors using a rolling-windows approach where each window is 10 years long. Similarly, I estimate the window-specific volatility and persistence of perceived inflation. Consistent with my theory of attention, I find strong negative correlations between inflation volatility and forecast errors, as well as between inflation persistence and forecast errors. This holds for both measures of

forecast errors, i.e., for the absolute values and the squared values of forecast errors. The correlations are indeed quite strong. For the squared forecast errors, I find a correlation with inflation persistence of -0.58 and with inflation volatility of -0.25. For the absolute values of forecast errors, the correlation with persistence is -0.64 and with inflation volatility -0.40. These results are consistent with the recent findings in Weber et al. (2023) who find that households that report to pay more attention to inflation have inflation expectations that are much closer to the actual level of inflation.

C   Model Details and Derivations

In this Appendix, I derive the New Keynesian Phillips Curve and the aggregate IS equation under limited attention (equations (7) and (8)). To nest the case of positive trend inflation (see Section 4.3.3), I do this for the general case that allows for an arbitrary steady state inflation rate, following Ascari and Rossi (2012) who derive the New Keynesian Phillips Curve with positive trend inflation with Rotemberg price adjustment costs (Rotemberg, 1982). A key assumption that I make throughout the following derivations is that firms set their prices optimally for a given inflation expectation, and these inflation expectations are provided by forecaster, as in Adam and Padula (2011).

Households.   There is a representative household obtaining utility from consumption and disutility from working, with lifetime utility

$$\tilde{E}_0 \sum_{t=0}^{\infty} \beta^t Z_t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Psi \frac{H_t^{1+\nu}}{1+\nu} \right], \tag{31}$$

where C_t is consumption of the final good, H_t is hours worked, β is the household's time discount factor, and $\tilde{E}_t$ denotes the household's subjective expectations operator based on information available in period t. Z_t are exogenous preference shocks. The parameters σ and ν pin down the relative risk aversion and the inverse Frisch labor elasticity, respectively. Ψ is the utility weight on hours worked.

Households maximize their lifetime utility subject to the flow budget constraints

$$C_t + B_t = w_t H_t + \frac{1+i_{t-1}}{1+\pi_t} B_{t-1} + \frac{T_t^H}{P_t}, \quad \text{for all } t, \tag{32}$$

where B_t is the real value of government bonds, w_t the real wage, π_t is the net inflation rate, and i_t the nominal interest rate. T_t^H denotes lump-sum taxes and transfers from the government.

Maximizing (31) subject to (32) yields the Euler equation

$$Z_t C_t^{-\sigma} = \beta(1+i_t)\tilde{E}_t \left[Z_{t+1} C_{t+1}^{-\sigma} \frac{1}{1+\pi_{t+1}} \right], \tag{33}$$

and the labor-leisure condition

$$w_t C_t^{-\sigma} = \Psi H_t^{\nu}. \tag{34}$$

Final goods producer. There is a representative final good producer that aggregates the intermediate goods $Y_t(j)$ to a final good Y_t, according to

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{35}$$

with $\epsilon > 1$. Nominal profits are given by $P_t \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}} - \int_0^1 P_t(j) Y_t(j) dj$, and profit maximization gives rise to the demand for each variety j:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t. \tag{36}$$

Thus, demand for variety j is a function of its relative price, the price elasticity of demand ϵ and aggregate output Y_t. The aggregate price level is given by

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}. \tag{37}$$

Intermediate goods producers. Intermediate producer of variety j produces output $Y_t(j)$ using labor $H_t(j)$ as its only input

$$Y_t(j) = H_t(j). \tag{38}$$

All intermediate producers pay the same wage w_t and a sales tax (or subsidy) τ_t, which in steady state is set such that profits in steady state are 0. These taxes are given back to firms in a lump-sum fashion, denoted $t_t^F(j)$. Taxes are assumed to be constant in the efficient economy, i.e., absent price rigidities, but fluctuate around their steady state in the economy with price rigidities in order to give rise to exogenous cost-push shocks.

Each intermediate firm has two managers: one is responsible for the firm's forecasts and the other manager sets the price of firm j given these forecasts, similar to the setup in, e.g., Adam and Padula (2011).

When adjusting the price, the firm is subject to a Rotemberg (1982) price-adjustment friction. Their per-period profits (in real terms) are given by

$$(1 - \tau_t) P_t(j) \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} \frac{Y_t}{P_t} - w_t H_t(j) - \frac{\psi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right)^2 Y_t + t_t^F(j), \tag{39}$$

where $\psi \geq 0$ captures the price-adjustment cost parameter. They set prices to maximize

$$Profits_0(j) = \tilde{E}_0^j \sum_{t=0}^{\infty} D_{0,t} \left[(1 - \tau_t) P_t(j) \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} \frac{Y_t}{P_t} - mc_t H_t(j) - \frac{\psi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right)^2 Y_t + t_t^F(j) \right],$$

52

where $D_{0,t} \equiv \beta^t \left(\frac{C_t}{C_0} \right)^{-\sigma}$ is the stochastic discount factor (for simplicity, I assume that firm managers are not subject to preference shocks), $mc_t = w_t$ denotes the real marginal cost which is the same for every firm. Using the production function to substitute for $H_t(j)$ and the demand for firm j's product from the final goods producer, the corresponding first order condition is then given by

$$T_t(\epsilon - 1) \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} = \epsilon mc_t \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon-1} - \psi \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right) \frac{P_t}{P_{t-1}(j)}$$
$$+ \beta \psi \tilde{E}_t^j \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \left(\frac{P_{t+1}(j)}{P_t(j)} - 1 \right) \frac{P_{t+1}(j)}{P_t(j)} \frac{P_t}{P_t(j)} \frac{Y_{t+1}}{Y_t} \right],$$

where $T_t \equiv 1 - \tau_t$.

Government. The government imposes a sales tax τ_t on sales of intermediate goods, issues nominal bonds, and pays lump-sum taxes and transfers T_t^H to households and $t_t^F(j)$ to firms. The real government budget constraint is given by

$$B_t = B_{t-1} \frac{1 + i_{t-1}}{\Pi_t} + \frac{T_t^H}{P_t} - \tau Y_t + t_t^f.$$

Lump-sum taxes and transfers are set such that they keep real government debt constant at the initial level B_{-1}/P_{-1}, which I set to zero.

Steady State. The resource constraint is given by $Y_t = C_t + \frac{\psi}{2} \left(\frac{P_t}{P_{t-1}} - 1 \right)^2 Y_t$. With steady state inflation denoted by $\bar{\Pi}$, it follows that in steady state we have

$$C = \left(1 - \frac{\psi}{2} (\bar{\Pi} - 1)^2 \right) Y. \tag{40}$$

From the production function, we have $Y = H$ and marginal costs are equal to the real wage, $mc = w$. Given the assumption that intermediate producers receive the subsidy τ which is set to induce the efficient steady state, it follows that $mc = w = 1$. Since all firms set the same price in steady state, it follows from the intermediate producers' first-order conditions, that

$$mc = \frac{(1 - \tau)(\epsilon - 1)}{\epsilon} + \psi \bar{\Pi}(\bar{\Pi} - 1) \frac{1}{\epsilon} (1 - \beta),$$

which implies that the steady state subsidy is equal to

$$T = 1 - \tau = \frac{\epsilon - \psi(\bar{\Pi} - 1)\bar{\Pi}(1 - \beta)}{\epsilon - 1}.$$

From the labor-leisure equation and the resource constraint, we obtain

$$Y = \left(\frac{1}{\Psi \left(1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2\right)^\sigma} \right)^{\frac{1}{\nu+\sigma}}.$$

Linearization. Linearizing the Euler equation (33) yields

$$\widehat{c}_t = \tilde{E}_t \widehat{c}_{t+1} - \varphi \left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1}) \right), \tag{41}$$

where $\varphi \equiv \frac{1}{\sigma}$. Linearizing the resource constraint, we obtain

$$\widehat{y}_t = \widehat{c}_t + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2} \pi_t,$$

where π_t now denotes inflation in deviations from its steady state value. Plugging this into (41), we get

$$\widehat{y}_t = \tilde{E}_t \widehat{y}_{t+1} + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2} \left[\pi_t - \tilde{E}_t \pi_{t+1} \right] - \varphi \left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1}) \right).$$

In order to express this in terms of the output gap, rather than output, we have to solve for the efficient output that prevails in the economy absent price rigidities (denoted by a "*"). From the production function, we have $Y_t^* = H_t^*$. The real wage is constant $w_t^* = 1$. From the labor-leisure equation (34), we get that potential output is therefore also constant and equal to

$$Y_t^* = \Psi^{-\frac{1}{\nu+\sigma}}. \tag{42}$$

Thus, potential output in log-deviations is 0. The Euler equation in the flexible-price economy is therefore given by

$$0 = -\varphi \left(r_t - (\widehat{z}_t - \tilde{E}_t z_{t+1}) \right).$$

Since the natural rate is defined as the real rate that prevails under flexible prices, r_t, it follows that

$$r_t^n = \widehat{z}_t - \tilde{E}_t z_{t+1}. \tag{43}$$

Substituting $\widehat{z}_t - \tilde{E}_t z_{t+1}$ with r_t^n in (41) and using that $\widehat{y}_t = y_t^{gap}$, since potential output in deviations from steady state is 0, yields the aggregate IS equation

$$y_t^{gap} = \tilde{E}_t y_{t+1}^{gap} + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2} \left[\pi_t - \tilde{E}_t \pi_{t+1} \right] - \varphi \left(i_t - \tilde{E}_t \pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1}) \right).$$

I assume for the most part of the analysis that output gap expectations are rational, $\tilde{E}_t y_{t+1}^{gap} = E_t y_{t+1}^{gap}$, and that households believe that inflation follows an AR(1) process and that they receive signals of the form $s_t = \pi_t + \varepsilon_t$ with normally distributed noise ε_t (see Section A for details). For tractability, I abstract from noise shocks and therefore assume that $\varepsilon_t = 0$ for all t but that the household behaves as if there was noise. This then gives rise to the law of motion for inflation expectations stated in equation (9). Taking everything together, we can therefore write

$$y_t^{gap} = E_t y_{t+1}^{gap} + \frac{\psi(\bar{\Pi}-1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi}-1)^2} \left[\pi_t - \pi_{t+1|t}^e\right] - \varphi\left(i_t - \pi_{t+1|t}^e - r_t^n\right).$$

In the case of zero trend inflation, $\bar{\Pi} = 1$, this collapses to

$$y_t^{gap} = E_t y_{t+1}^{gap} - \varphi\left(i_t - \pi_{t+1|t}^e - r_t^n\right),$$

as stated in equation (8).

In order to derive the Phillips Curve, we need to linearize the intermediate producers' first-order condition. This condition is given by

$$\underbrace{T_t(\epsilon-1)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}}_{I} = \underbrace{\epsilon mc_t \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon-1}}_{II} \underbrace{-\psi\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)\frac{P_t}{P_{t-1}(j)}}_{III}$$

$$\underbrace{+ \beta\psi\tilde{E}_t^j \left[\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\left(\frac{P_{t+1}(j)}{P_t(j)} - 1\right)\frac{P_{t+1}(j)}{P_t(j)}\frac{P_t}{P_t(j)}\frac{Y_{t+1}}{Y_t}\right]}_{IV}$$

The linearization of the terms I to IV yields

$$I: \quad (\epsilon-1)T\widehat{T}_t - \epsilon(\epsilon-1)T\widehat{p}_t^j + (\epsilon-1)\epsilon T\widehat{p}_t$$

$$II: \quad \epsilon(\widehat{mc}_t - (1+\epsilon)\widehat{p}_t^j + (1+\epsilon)\widehat{p}_t)$$

$$III: \quad -\psi\bar{\Pi}(\bar{\Pi}-1)\widehat{p}_t - \psi\bar{\Pi}^2\widehat{p}_t^j + \psi\bar{\Pi}(2\bar{\Pi}-1)\widehat{p}_{t-1}^j$$

$$IV: \quad \beta\psi\bar{\Pi}^2\tilde{E}_t^j\pi_{t+1}^j + \beta\psi\bar{\Pi}(\bar{\Pi}-1)\left[(1-\sigma)\tilde{E}_t^j\widehat{y}_{t+1} + (\sigma-1)\widehat{y}_t\right.$$

$$\left. + \frac{\sigma\psi(\bar{\Pi}-1)}{1 - \frac{\psi}{2}(\bar{\Pi}-1)^2}\left(\tilde{E}_t^j\pi_{t+1} - \pi_t\right) + \tilde{E}_t^j\pi_{t+1}^j + \widehat{p}_t - \widehat{p}_t^j\right],$$

where I used the linearized resource constraint to arrive at the expression IV. When trend inflation is zero, $\bar{\Pi} = 1$, the expression IV becomes simply $\beta\psi\tilde{E}_t^j\pi_{t+1}^j$. Since, I focus on the case $\sigma = 1$, the terms relating to output in expression IV drop out. Thus, the only reason why prices may differ across firms j is due to different forecasts of future inflation (either of

aggregate or of firm-specific inflation). Following the assumption in Adam and Padula (2011), I assume that these forecasts are provided by forecasters that are different from the price setting managers. Given that there are no idiosyncratic shocks, I assume that the forecaster of firm j expects firm-specific inflation to be equal to aggregate inflation, $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t^j \pi_{t+1}$. These forecasters then form their inflation expectations in the same way as households and as detailed in the limited attention problem in Section 2. Given that I abstract from noise (but signals are perceived as noisy), all forecasters then receive exactly the same signals and therefore also form their expectations equally. Note, that a weaker assumption would be sufficient to arrive at the following results: namely, that all forecasters receive the same signal.[28] Thus, $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t^j \pi_{t+1} = \tilde{E}_t \pi_{t+1}$. Therefore, all price setters set the same price, $\hat{p}_t^j = \hat{p}_t$. Thus, the belief that firm-specific inflation coincides with aggregate inflation is satisfied in equilibrium, which confirms the forecasters' belief and she therefore does not have an incentive to update that belief.

Using this, and

$$\widehat{mc}_t = (1+\nu)\widehat{y}_t - \frac{\psi(\bar{\Pi}-1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi}-1)^2}\pi_t,$$

which follows from the labor-leisure equation, the production function and the resource constraint, we then obtain the following Phillips Curve with trend inflation and limited attention:

$$\pi_t = \zeta \left[\frac{\epsilon(1+\nu)}{\psi} \widehat{y}_t^{gap} + \beta\bar{\Pi}^2 \pi_{t+1|t}^e + u_t + \Xi \pi_{t+1|t}^e \right]$$

where

$$\zeta \equiv \frac{1}{\bar{\Pi}(2\bar{\Pi}-1) + \frac{\beta\psi\bar{\Pi}(\bar{\Pi}-1)^2}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2} + \frac{\epsilon\bar{\Pi}(\bar{\Pi}-1)}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}}$$

$$\Xi \equiv \beta\bar{\Pi}(\bar{\Pi}-1)\left[1 + \frac{\psi(\bar{\Pi}-1)}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}\right],$$

as stated in Section 4.3.3, and with the cost-push shock defined as $u_t \equiv -\frac{(\epsilon-1)T}{\psi}\widehat{T}_t$. For the case of zero trend inflation, we get $\zeta = 1$ and $\Xi = 0$, so that the Phillips Curve reduces to

$$\pi_t = \beta\pi_{t+1|}^e + \kappa y_t^{gap} + u_t,$$

which is equation (7).

---

[28]If all forecasters receive the same signal, they form identical inflation expectations. However, the noise term would give rise to a "belief shock" in the Phillips Curve (and also the aggregate IS equation). I ignore these belief shocks.

D   Analytical Results and Proofs

To see how lower attention weakens the effectiveness of forward guidance, consider the following stylized experiment.[29] The economy is hit by a negative natural rate shock in period $t = 0$ that pushes the nominal interest rate to the effective lower bound, i.e., $r_0^n < 0$ and $i_0 = -\underline{i}$. In $t = 1$, the natural rate returns to its steady state value and stays there indefinitely, $r_t^n = 0$ for all $t \geq 1$. From period $t = 2$ onwards, the output gap, and the real rate are back at their steady states, $y_t^{gap} = 0$ and $i_t - \pi_{t+1|t}^e = 0$ for all $t \geq 2$.

To model forward guidance, the real rate is assumed to be below the natural rate in $t = 1$. To make it comparable across different degrees of attention, I impose that

$$r_1 \equiv i_1 - \pi_{2|1}^e < 0$$

is the same for all γ and known in advance.[30] Hence, forward guidance here means to announce a certain value for the real rate. I discuss the implications of forward guidance via the nominal rate in section D.1. In the following, I assume that $(-\underline{i} - r_0^n + r_1)$ is negative, which means that the announced policy, captured by $r_1 < 0$, *makes up* for the binding lower bound in $t = 0$, captured by $-\underline{i} - r_0^n > 0$.

Given the real rate r_1 and the fact that $y_2^{gap} = 0$, the Euler equation in $t = 1$ determines the output gap in period 1 as

$$y_1^{gap} = -\varphi(r_1) > 0. \tag{44}$$

Equation (44) captures the *make-up policy:* by keeping the real rate below the natural rate, output is above potential after the lower-bound constraint stops to be binding.

In $t = 0$, the ELB binds and the natural rate is negative. Thus, the Euler equation in $t = 0$ yields

$$y_0^{gap} = \underbrace{-\varphi(r_1)}_{=E_0 y_1^{gap}} - \varphi\left(-\underline{i} - \pi_{1|0}^e - r_0^n\right).$$

Substituting the law of motion for inflation expectations

$$\pi_{1|0}^e = (1-\gamma)\pi_{0|-1}^e + \gamma\pi_0,$$

into the Phillips Curve

$$\pi_0 = \frac{\beta}{1-\beta\gamma}(1-\gamma)\pi_{0|-1}^e + \frac{\kappa}{1-\beta\gamma}y_0^{gap}$$

---

yields an expression for inflation expectations:

$$\pi^e_{1|0} = \frac{1-\gamma}{1-\beta\gamma}\pi^e_{0|-1} + \frac{\kappa\gamma}{1-\beta\gamma}y^{gap}_0.$$

Putting everything together, we arrive at the following result.

Proposition 2. *The output gap in the period when the shock hits, $t = 0$, is given by*

$$y^{gap}_0 = -\frac{\varphi(1-\beta\gamma)}{1-\gamma(\beta+\varphi\kappa)}[-\underline{i} - r^n_0 + r_1] + \frac{\varphi(1-\gamma)}{1-\gamma(\beta+\varphi\kappa)}\pi^e_{0|-1} \tag{45}$$

and inflation in $t = 0$ is given by

$$\pi_0 = -\frac{\kappa\varphi}{1-\gamma(\beta+\varphi\kappa)}[-\underline{i} - r^n_0 + r_1] + (1-\gamma)\left[\frac{\beta}{1-\beta\gamma} + \frac{\varphi}{1-\gamma(\beta+\varphi\kappa)}\right]\pi^e_{0|-1}. \tag{46}$$

Proposition 2 captures the effectiveness of forward guidance on the output gap and inflation in the period when the shock hits. Assuming $(1-\gamma(\beta+\varphi\kappa))$ is positive makes sure that forward guidance has a stimulating effect on output and inflation in $t = 0$. Proposition 2 captures several channels how a change in attention affects the economy's response to forward guidance, which I discuss in the following two corollaries.

Corollary 2. *Lower attention weakens*

  (i) *the negative effects of the shock,*

 (ii) *the positive effects of forward guidance,*

(iii) *the positive effects of a decrease in the lower bound $-\underline{i}$*

on the output gap and inflation.

Corollary 2 follows from the fact that the terms $\frac{\varphi(1-\beta\gamma)}{1-\gamma(\beta+\varphi\kappa)}$ and $\frac{\kappa\varphi}{1-\gamma(\beta+\varphi\kappa)}$ in front of $[-\underline{i} - r^n_0 + r_1]$ are both increasing in γ. Points (i) and (ii) capture the main trade off of lower attention. While lower attention has a stabilizing effect via more anchored inflation expectations (point (i)), it renders forward guidance less effective (point (ii)). The reason why forward guidance becomes less effective as attention declines is because inflation expectations increase less in response to the announced policy, and thus, the real rate remains higher. Point (iii) illustrates an additional drawback of lower attention. A reduction of the effective lower bound, $-\underline{i}$, is less stimulating if agents in the economy are less attentive. Thus, going from a zero lower bound to a lower bound in negative territory, as conducted in several advanced economies over the last ten years, becomes less effective in terms of stimulating output and inflation if the public is inattentive (consistent with the exercise in figure 5). Away from the lower bound, point (iii) implies that

the effectiveness of conventional monetary policy via the nominal interest rate becomes less effective as attention declines.

How attention matters for the transmission of prior inflation expectations on the output gap and inflation is ambiguous, as the following Corollary shows.

Corollary 3. *Lower attention*

(i) *weakens the positive effect of higher prior inflation beliefs, $\pi^e_{0|-1}$, on the output gap if and only if,*

$$(\beta + \varphi\kappa) > 1, \tag{47}$$

(ii) *weakens the positive effect of higher prior inflation beliefs on inflation if and only if*

$$\frac{\beta(\beta-1)}{(1-\beta\gamma)^2} + \frac{\varphi((\beta+\varphi\kappa)-1)}{(1-\gamma(\beta+\varphi\kappa))^2} > 0. \tag{48}$$

Overall, the role of attention for the effects of higher prior beliefs on output and inflation is ambiguous. This is mainly the case because, on the one hand, lower attention implies that agents put more weight on their prior beliefs. On the other hand, as discussed previously, lower attention leads to more stable inflation overall, thus, weakening the effects of prior beliefs.

Given the calibration in Table 4, conditions (47) and (48) both hold for all $\gamma < 0.99$. The effects of changes in γ, however, are numerically small. Thus, an increase in the average inflation rate—which increases average prior beliefs—is a promising monetary instrument to combat the loss of control via forward guidance as attention declines. By *ex-ante* increasing the average inflation rate, the policymaker not only supports higher inflation expectations and thus, lower real rates for a given nominal rate, but also gains additional policy space through the increase in the average nominal rate.

D.1 Extensions

I now show that all the results go through when relaxing the assumption that $\rho_\pi = 1$ and also discuss how forward guidance via the nominal (instead of the real) interest rate changes the results and I also allow for attention heterogeneity across firms and households. We consider the same stylized experiment but now the law of motion for inflation expectations is given by

$$\pi^e_{1|0} = (1 - \rho_\pi)\bar{\pi} + \rho_\pi(1-\gamma)\pi^e_{0|-1} + \rho_\pi\gamma\pi_0,$$

which can be substituted into the Phillips Curve:

$$\pi_0 = \frac{\beta}{1 - \beta\rho_\pi\gamma} \left((1 - \rho_\pi)\bar{\pi} + \rho_\pi(1-\gamma)\pi^e_{0|-1}\right) + \frac{\kappa}{1 - \beta\rho_\pi\gamma} y^{gap}_0.$$

Thus, inflation expectations are given by

$$\pi_{1|0}^e = \frac{1 - \rho_\pi}{1 - \beta\rho_\pi\gamma}\bar{\pi} + \frac{\rho_\pi(1 - \gamma)}{1 - \beta\rho_\pi\gamma}\pi_{0|-1}^e + \frac{\kappa\rho_\pi\gamma}{1 - \beta\rho_\pi\gamma}y_0^{gap}.$$

Putting everything together, we arrive at the following Proposition.

Proposition 3. *The output gap in the period when the shock hits, $t = 0$, is given by*

$$y_0^{gap} = -\frac{\varphi(1 - \beta\rho_\pi\gamma)}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}[-\underline{i} - r_0^n + r_1]$$

$$+ \frac{\varphi}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}\left[(1 - \rho_\pi)\bar{\pi} + \rho_\pi(1 - \gamma)\pi_{0|-1}^e\right]$$

and inflation in $t = 0$ is given by

$$\pi_0 = -\frac{\kappa\varphi}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}[-\underline{i} - r_0^n + r_1] + (1 - \rho_\pi)\left[\frac{\beta}{1 - \beta\rho_\pi\gamma} + \frac{\varphi}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}\right]\bar{\pi}$$

$$+ \rho_\pi(1 - \gamma)\left[\frac{\beta}{1 - \beta\rho_\pi\gamma} + \frac{\varphi}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}\right]\pi_{0|-1}^e. \tag{49}$$

Proposition 3 captures the effectiveness of forward guidance on the output gap and inflation in the period when the shock hits. The assumption that $(1 - \rho_\pi\gamma(\beta + \varphi\kappa))$ is positive, makes sure that forward guidance, i.e, a lower r_1 has a stimulating effect on output and inflation in $t = 0$. Proposition 3 captures several channels how a change in attention affects the economy's response to forward guidance, which I now collect in a series of corollaries.

Corollary 4. *Lower attention*

(i) *weakens the negative effect of the shock on impact,*

(ii) *weakens the effects of forward guidance on the output gap and inflation,*

(iii) *weakens the stimulative effects of a decrease in the lower bound $-\underline{i}$.*

Corollary 4 follows from the fact that the terms $\frac{\varphi(1 - \beta\rho_\pi\gamma)}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}$ and $\frac{\kappa\varphi}{1 - \rho_\pi\gamma(\beta + \varphi\kappa)}$ in front of $[-\underline{i} - r_0^n + r_1]$ are both increasing in γ. Points (i) and (ii) capture the main trade off of lower attention. While lower attention has a stabilizing effect via more anchored inflation expectations (point (i)), it renders forward guidance less effective (point (ii)). The reason why forward guidance becomes less effective as attention declines is because inflation expectations increase less in response to the announced policy, and thus, the real rate remains higher. Point (iii) illustrates an additional drawback of lower attention. A reduction of the effective lower bound, $-\underline{i}$, is less stimulating if agents in the economy are less attentive. Thus, going from a zero lower bound to a lower bound in negative territory, as conducted in several advanced

economies over the last ten years, becomes less effective in terms of stimulating output and inflation if the public is inattentive. Note, that a decrease in the perceived inflation persistence, ρ_π, has the exact same implications as a decrease in γ.

The next corollary discusses how changes in attention affect the role of long-run inflation beliefs on the output gap and inflation.

Corollary 5. *Lower attention weakens the positive effects of higher long-run inflation beliefs $\bar{\pi}$ on output and inflation,*

Corollary 5 says that higher long-run beliefs have a positive effect on inflation and the output gap, but lower attention weakens these effects. However, as long as $\gamma(\beta + \varphi\kappa) < 1$, a higher ρ_π mutes the effects of $\bar{\pi}$ on the output gap. Since this condition is usually satisfied and because ρ_π is in general close to 1, the role of high long-run inflation beliefs is quite weak. In the limit case $\rho_\pi \to 1$, long-run beliefs become irrelevant.

How attention matters for the transmission of prior inflation expectations on the output gap and inflation is ambiguous, as the following Corollary shows.

Corollary 6. *Lower attention*

(i) *weakens the positive effect of higher prior inflation beliefs, $\pi^e_{0|-1}$, on the output gap if and only if,*

$$\rho_\pi (\beta + \varphi\kappa) > 1, \tag{50}$$

(ii) *weakens the positive effect of higher prior inflation beliefs on inflation if and only if*

$$\frac{\rho_\pi \beta (\rho_\pi \beta - 1)}{(1 - \beta\rho_\pi\gamma)^2} + \frac{\rho_\pi \varphi (\rho_\pi(\beta + \varphi\kappa) - 1)}{(1 - \rho_\pi\gamma(\beta + \varphi\kappa))^2} > 0. \tag{51}$$

Overall, the role of attention for the effects of higher prior beliefs on output and inflation is ambiguous. This is mainly the case because, on the one hand, lower attention implies that agents put more weight on their prior beliefs. On the other hand, as discussed previously, lower attention leads to more stable inflation overall, thus, weakening the effects of prior beliefs. This can also be seen in the discussion of the Phillips Curve, see Proposition 1.

Given the calibration in Table 4, conditions (50) and (51) both hold. The effects of changes in γ, however, are numerically small. Thus, an increase in the average inflation rate—which also increases average prior beliefs—is a promising monetary instrument to combat the loss of control via forward guidance as attention declines. By *ex-ante* increasing the average inflation rate, the policymaker not only supports higher inflation expectations and thus, lower real rates for a given nominal rate, but also gains additional policy space through the increase in the average nominal rate. Higher average inflation, however, is also costly. In the analysis of optimal policy,

later on, I will explore this trade off and characterize the optimal inflation target for different levels of attention.

D.1.1 Forward Guidance via Nominal Interest Rates

So far, forward guidance was characterized as a promise to keep the *real* rate low. Now, assume that forward guidance is conducted via promising lower *nominal* rates instead. Thus, i_1 will be fixed across different γ. For simplicity, I focus on the case with $\rho_\pi = 1$ and $\pi^e_{0|-1} = 0$. It follows from the Euler equation in $t = 1$ that

$$y_1^{gap} = -\varphi \left(i_1 - (1 - \gamma)\gamma\pi_0 - \gamma\pi_1 \right).$$

The Phillips Curve in $t = 1$ yields

$$\pi_1 = \frac{(1 - \gamma)\gamma}{1 - \beta\gamma}\pi_0 + \frac{\kappa}{1 - \beta\gamma}y_1^{gap},$$

so that we get an expression for y_1^{gap} in terms of π_0:

$$y_1^{gap} = -\frac{\varphi(1 - \beta\gamma)}{1 - \gamma(\beta + \varphi\kappa)}i_1 + \varphi(1 - \gamma)\gamma\frac{1 + \gamma(1 - \beta)}{1 - \gamma(\beta + \varphi\kappa)}\pi_0. \tag{52}$$

Given $\pi^e_{1|0} = \gamma\pi_0$, the Phillips Curve in $t = 0$ yields

$$\pi_0 = \frac{\kappa}{1 - \beta\gamma}y_0^{gap},$$

and hence, $\pi^e_{1|0} = \frac{\kappa\gamma}{1 - \beta\gamma}y_0^{gap}$. Plugging this into the Euler equation in $t = 0$ gives

$$y_0^{gap} = \mathbb{E}_0 y_1^{gap} - \varphi\left(-\underline{i} - \frac{\kappa\gamma}{1 - \beta\gamma}y_0^{gap} - r_0^n\right).$$

Solving for y_0^{gap} leads to the following Lemma.

Lemma 7. *Forward guidance via the nominal interest rate yields the following output gap*

$$y_0^{gap} = A_1\left[-\frac{\varphi(1 - \beta\gamma)}{1 - \gamma(\beta + \varphi\kappa)}i_1 - \varphi\left(-\underline{i} - r_0^n\right)\right], \tag{53}$$

and inflation

$$\pi_0 = \frac{\kappa}{1 - \beta\gamma}A_1\left[-\frac{\varphi(1 - \beta\gamma)}{1 - \gamma(\beta + \varphi\kappa)}i_1 - \varphi\left(-\underline{i} - r_0^n\right)\right], \tag{54}$$

where

$$A_1 \equiv \frac{1}{1 - \varphi(1-\gamma)\gamma\frac{1+\gamma(1-\beta)}{1-\gamma(\beta+\varphi\kappa)}\frac{\kappa}{1-\beta\gamma} - \frac{\varphi\kappa\gamma}{1-\beta\gamma}}. \tag{55}$$

Given the calibration in Table 4, A_1 is positive and increasing in γ. Thus, promising lower future nominal interest rates can indeed stimulate the economy. But similar to the case in which the policy maker commits to a certain future *real* rate, forward guidance becomes less effective when agents are less attentive. In fact, all three results from Corollary 4 go through.

Recall equation (52):

$$y_1^{gap} = -\frac{\varphi(1-\beta\gamma)}{1 - \gamma(\beta+\varphi\kappa)}i_1 + \varphi(1-\gamma)\gamma\frac{1+\gamma(1-\beta)}{1-\gamma(\beta+\varphi\kappa)}\pi_0. $$

Note, that the first term becomes less negative as γ declines. Given the calibration in Table 4, also the second term decreases as attention declines. Thus, for a given π_0, a particular i_1 has weaker effects on the output gap in $t = 1$ at lower levels of attention. Since lower attention also weakens the positive effects of forward guidance on π_0, the output gap (and inflation) stay lower also in $t = 1$.

Since inflation in $t = 0$ and $t = 1$ is lower at smaller values of γ, also $\pi_{2|1}^e$ will be lower and thus, for a given nominal rate i_1, the *real* rate, $r_1 \equiv i_1 - \pi_{2|1}^e$, will be higher. Hence, to achieve a certain forward guidance in terms of the real interest rate, the promise in terms of the nominal rate needs to be larger when firms and households are inattentive. Combining this with the findings on the effectiveness of forward guidance via the *real* rate (Proposition 3) shows how lower attention renders forward guidance less powerful *even though* the promise in terms of the *nominal rate* is stronger.

D.1.2 Heterogeneous Attention

So far, I assumed that firms and households are equally attentive. But what if firms and households differ in their attention to inflation? Let us denote firms' attention by γ_F and households' attention by γ_H with $\gamma_F \neq \gamma_H$. For clarity, I focus on the case with $\rho_\pi = 1$ and $\pi_{0|-1}^{e,j} = 0$ for $j \in \{F, H\}$.

Lemma 8. *With heterogeneous attention to inflation, the output gap in $t = 0$ is given by*

$$y_0^{gap} = \frac{-\varphi(1-\beta\gamma_F)}{1 - \beta\gamma_F - \kappa\varphi\gamma_H}[-\underline{i} + r_1 - r_0^n], \tag{56}$$

and inflation by

$$\pi_0 = \frac{-\varphi\kappa}{1 - \beta\gamma_F - \kappa\varphi\gamma_H}[-\underline{i} + r_1 - r_0^n], \tag{57}$$

where $r_1 \equiv i_1 - \pi_{2|1}^{e,H}$ is the real rate given the households' expectations.

Lemma 8 shows that a similar result as in Corollary 4 holds under heterogeneous attention levels.

Corollary 9. *Lower attention of either firms or households*

   (i) weakens the negative effect of the shock on the output gap and inflation on impact,

   (ii) weakens the effects of forward guidance on the output gap and inflation,

   (iii) weakens the stimulative effects of a decrease in the lower bound $-\underline{i}$ on the output gap and inflation.

The parts concerning the output gap in Corollary 9 follow because the term in front of the brackets in equation (56) becomes more negative as either of $\{\gamma_F, \gamma_H\}$ increases:

$$\frac{\partial \left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H} \right]}{\partial \gamma_F} = -\frac{\beta\kappa\varphi^2\gamma_H}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^2} < 0$$

$$\frac{\partial \left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H} \right]}{\partial \gamma_H} = -\frac{\varphi^2(1-\beta\gamma_F)}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^2} < 0,$$

and the parts concerning inflation because the term $\frac{-\varphi\kappa}{1-\beta\gamma_F-\kappa\varphi\gamma_H}$ in equation (57) becomes more negative as either of $\{\gamma_F, \gamma_H\}$ increases, too.

Thus, if either firms or households (or both) become less attentive, forward guidance becomes less effective. In fact, the two degrees of attention reinforce each other, as the following Corollary shows.

Corollary 10. *Lower levels of households' attention to inflation weaken the effectiveness of forward guidance, especially when firms' attention to inflation is low, and vice-versa.*

To see this, note that

$$\frac{\partial^2 \left[-\frac{\varphi\kappa}{1-\beta\gamma_F-\kappa\varphi\gamma_H} \right]}{\partial \gamma_F \partial \gamma_H} = \frac{-2\varphi^2\kappa^2\beta}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^3} < 0,$$

$$\frac{\partial^2 \left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H} \right]}{\partial \gamma_F \partial \gamma_H} = \frac{-\beta\kappa\varphi^2 \left[1-\beta\gamma_F+\kappa\varphi\gamma_H \right]}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^3} < 0.$$

64

D.2 Proof of Proposition 1

Proof. The New Keynesian Phillips Curve is given by

$$\pi_t = \beta \pi_{t+1|t}^e + \kappa y_t^{gap} + u_t.$$

Substituting

$$\pi_{t+1|t}^e = \pi_{t|t-1}^e + \gamma \left(\pi_t - \pi_{t|t-1}^e \right)$$

for $\pi_{t+1|t}^e$ yields

$$
\begin{aligned}
\pi_t &= \beta \left(\pi_{t|t-1}^e + \gamma \left(\pi_t - \pi_{t|t-1}^e \right) \right) + \kappa y_t^{gap} + u_t \\
\Leftrightarrow & \pi_t (1 - \beta\gamma) = \beta \pi_{t|t-1}^e (1 - \gamma) + \kappa y_t^{gap} + u_t \\
\Leftrightarrow & \pi_t = \frac{\beta \pi_{t|t-1}^e (1 - \gamma) + \kappa y_t^{gap} + u_t}{(1 - \beta\gamma)} \\
\Leftrightarrow & \pi_t = \frac{\beta (1 - \gamma)}{(1 - \beta\gamma)} \pi_{t|t-1}^e + \frac{\kappa}{(1 - \beta\gamma)} y_t^{gap} + \frac{u_t}{(1 - \beta\gamma)}.
\end{aligned}
$$

Now, taking derivatives with respect to y_t^{gap}, u_t, and $\pi_{t|t-1}^e$, respectively, yields the results (i), (ii), and (iii). □

E   Additional Numerical Results

E.1   Non-Rational Output Gap Expectations

In this section, I estimate households' attention to the output gap (using expected unemployment changes as a proxy) and derive the policy implications of non-rational output gap expectations. I use the Survey of Consumers from the University of Michigan which asks respondents about what they think will happen to unemployment over the next 12 months. A drawback of that question is that respondents give a qualitative answer, saying that they expect unemployment to either "go up", "stay about the same", or "go down". Following Bhandari et al. (2022), I translate these qualitative answers into quantitative answers (see Bhandari et al. (2022), or Pfäuti and Seyrich (2022) for details). One assumption I need for this is that I have to impose what "about the same" means. I assume that survey respondents answer "about the same" when they believe that unemployment will change less than 0.15% which is half a standard deviation of unemployment changes over the period 1978-2019.

I then estimate attention to unemployment, γ^y, in the same way I estimate attention to inflation in Section 2, and I do so separately for the period before 1990 and the period after 1990. Table 20 shows the results. Attention to unemployment slightly increased from 0.088 before the 1990s ($\widehat{\gamma}_{y,<1990}$) to 0.100 after the 1990s ($\widehat{\gamma}_{y,\geq1990}$). These differences, however, are not statistically significant, as the last column indicates. Similarly, when I set the break point at 2000, I estimate attention levels of 0.098 before 2000, and 0.099 after 2000. Again, the difference between the two is not statistically significantly different from 0. These results therefore indicate that while there was a strong decline in people's attention to inflation, their attention to unemployment did not change.

Table 20: Attention to Unemployment

	$\widehat{\gamma}_{y,<1990}$	$\widehat{\gamma}_{y,\geq1990}$	p-val. $\widehat{\gamma}_{y,<1990} = \widehat{\gamma}_{y,\geq1990}$
Estimate	0.088	0.100	0.554
s.e.	(0.0379)	(0.0263)	

Notes: This table shows the estimated attention parameters with respect to unemployment, separately for the period before the 1990s ($\widehat{\gamma}_{y,<1990}$) and the period after 1990 ($\widehat{\gamma}_{y,\geq1990}$). The last column shows the p-value for the null-hypothesis that the two coefficients are the same. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey-West with four lags).

Now, to understand the policy implications of limited attention to the output gap, I impose that output gap expectations are given by

$$y_{t+1|t}^{gap,e} = y_{t|t-1}^{gap,e} + \gamma^y \left(y_t^{gap} - y_{t|t-1}^{gap,e} \right). \tag{58}$$

With these expectations, the aggregate IS equation is given by

$$y_t^{gap} = y_{t+1|t}^{gap,e} - \varphi \left(i_t - \pi_{t+1|t}^e - r_t^n \right), \tag{59}$$

whereas the Phillips Curve and the Taylor rule remain unchanged.

Figure 10 shows the impulse response functions of the main variables in this economy after a negative three-standard deviation natural rate shock and with $\gamma^y = 0.1$ (I set γ^π to 0.3, as in Section 3 and keep the rest of the calibraton also unchanged). We see that the inflation-attention traps get exacerbated. The reason is that now make-up policies are even less effective because not only inflation expectations are backward looking but also output gap expectations. Thus, even though there is interest-rate smoothing in the Taylor rule which features some form of make-up policy, this is not effective in stimulating expectations and thus, the economy remains stuck at the ELB even longer. Furthermore, inflation, inflation expectations, and now also the output gap stay below their initial values very persistently.

Table 21 shows the implications of limited attention to the output gap for optimal policy. The upper part of the table shows the optimal inflation target and welfare when output gap expectations are rational, and the lower part shows the results when output gap expectations are non-rational. The table highlights the following two main results: (i) limited attention to the output gap increases the optimal inflation target quite substantially and decreases welfare (independent of the level of γ^y or γ^π), and (ii) higher attention to the output gap reduces the optimal inflation target and increases welfare. The second result mirrors the main result regarding attention to inflation: lower attention (to inflation or the output gap) is welfare deteriorating in the presence of an effective lower bound constraint on nominal interest rates.

Table 21: Non-rational output gap expectations

	Inflation Target	Welfare
Rational $E_t y_{t+1}^{gap}$		
$\gamma^\pi = 0.25$	1.18%	-0.0053
$\gamma^\pi = 0.2$	1.20%	-0.0054
$\gamma^\pi = 0.1$	1.82%	-0.0105
Limited attention $y_{t+1|t}^{gap,e}$		
$\gamma^\pi = 0.25, \gamma^y = 0.075$	2.46%	-0.009
$\gamma^\pi = 0.2, \gamma^y = 0.075$	3.22%	-0.014
$\gamma^\pi = 0.2, \gamma^y = 0.125$	2.99%	-0.012
$\gamma^\pi = 0.1, \gamma^y = 0.125$	3.04%	-0.013
$\gamma^\pi = 0.1, \gamma^y = 0.15$	2.86%	-0.011

Notes: This table shows the implications of limited attention to the output gap for the optimal inflation target and welfare, for different combinations of γ^y (attention to the output gap) and γ^π (attention to inflation).

Figure 10: Impulse Response Functions with Non-Rational Output Gap Expectations
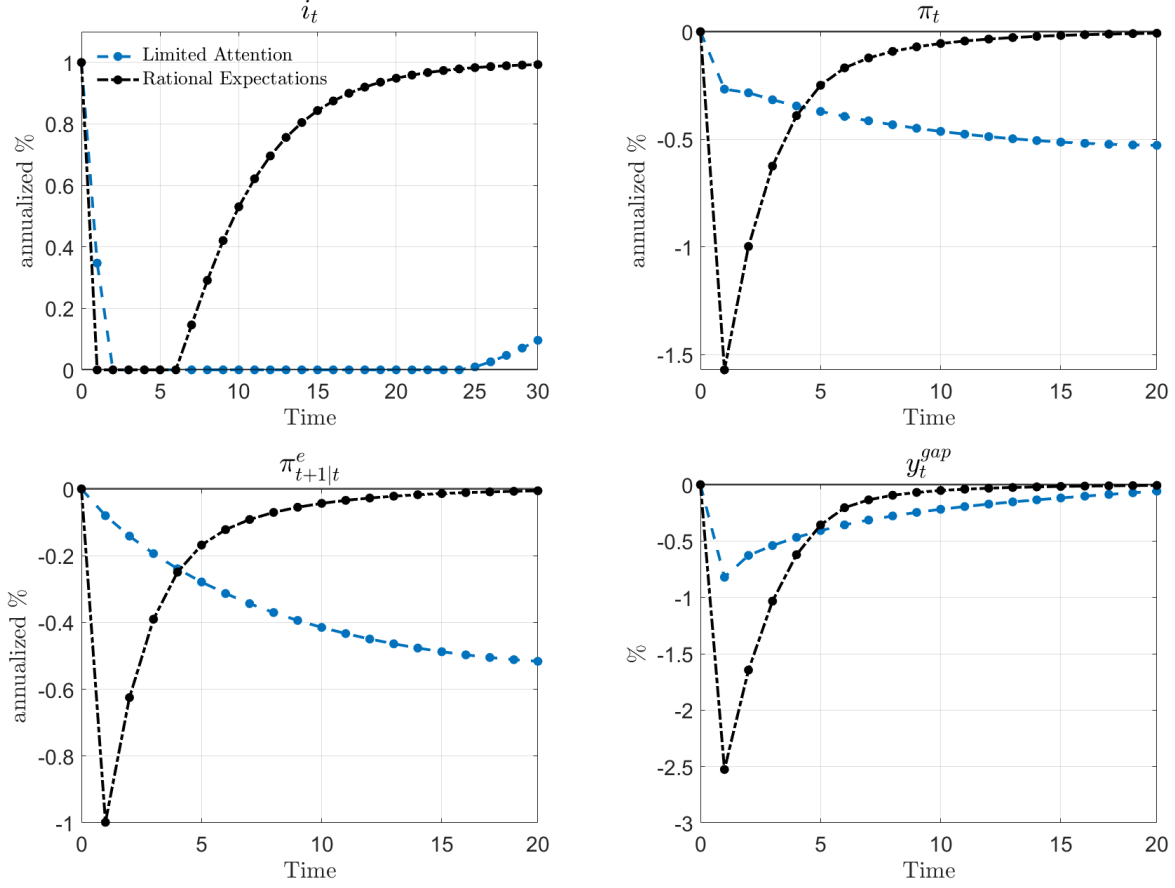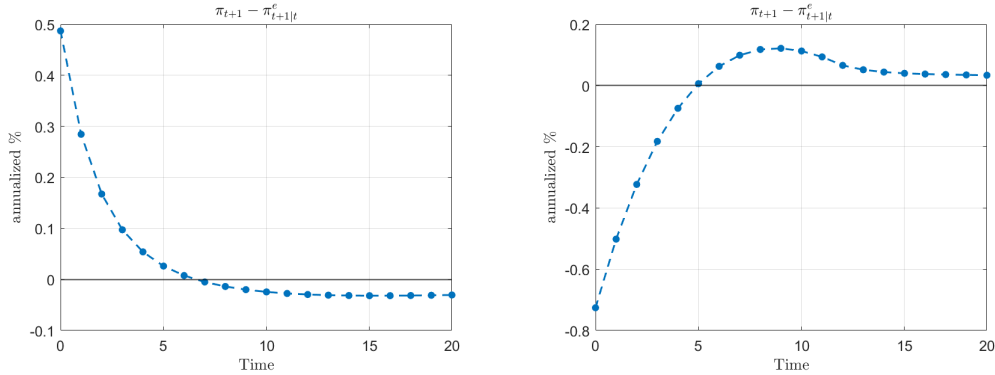

Note: This figure shows the impulse-response functions of the nominal interest rate (upper-left panel), inflation (upper-right panel), inflation expectations (lower-left) and the output gap (lower-right) to a negative natural rate shock of three standard deviations in the case where output gap expectations are given by equation (58). The blue-dashed lines show the case for the limited-attention model and the black-dashed-dotted lines for the rational expectations model. Everything is in terms of percentage deviations from the respective steady state levels, except the nominal rate is in levels.

E.2   Different Taylor Rule

To show that the exact specification of the Taylor rule is not essential for the occurence of inflation-attention traps, Figure 11 shows the impulse-response functions of the nominal interest rate, inflation, inflation expectations and the output gap for the model in which the Taylor rule absent the ELB is given by

$$i_t = 1.5\pi_t. \tag{60}$$

Figure 11: IRFs to Natural Rate Shock for Taylor rule (60)
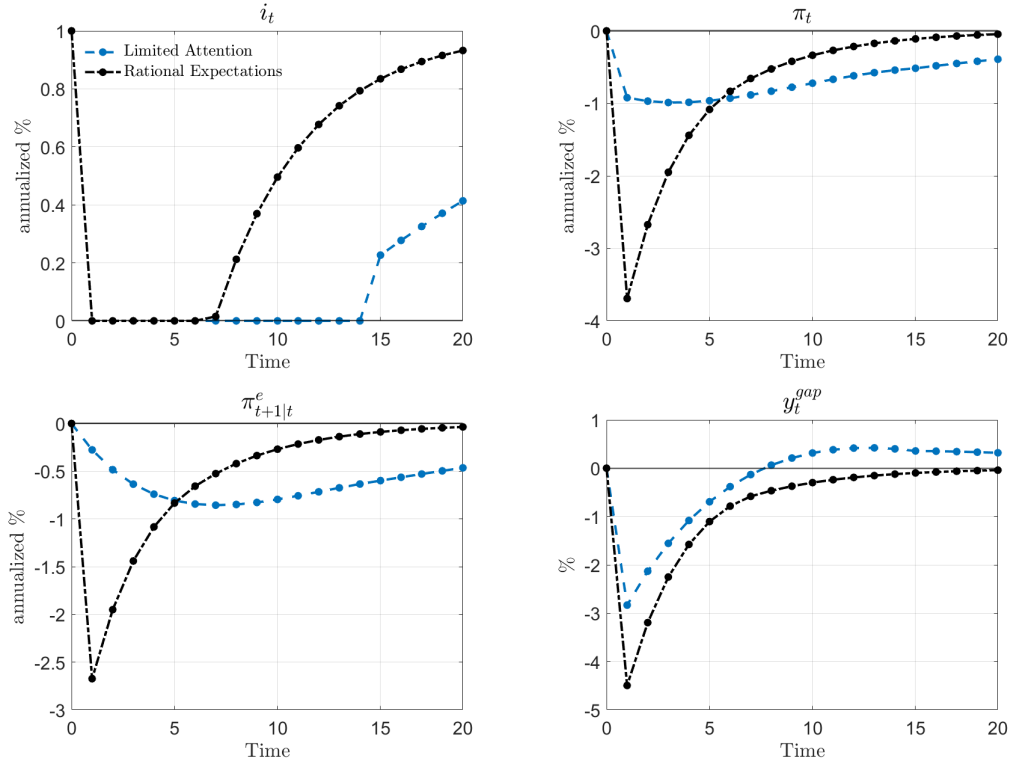
Note: This figure shows the impulse-response functions of the nominal interest rate (upper-left panel), inflation (upper-right panel), inflation expectations (lower-left) and the output gap (lower-right) to a negative natural rate shock of three standard deviations. The blue-dashed lines show the case for the limited-attention model and the black-dashed-dotted lines for the rational expectations model. Everything is in terms of percentage deviations from the respective steady state levels, except the nominal rate is in levels.

E.3 Forecast Errors

Angeletos et al. (2021) propose a new test of models that deviate from FIRE. Namely, that expectations should initially underreact but overshoot eventually. A straightforward way to test this is to look at the model-implied impulse response functions of the forecast error, $\pi_{t+1} - \pi^e_{t+1|t}$, to an exogenous shock. Figure 12 shows these IRFs. The left panel shows the IRF of the forecast error after a positive natural rate shock and the right panel shows the corresponding IRF to a negative natural rate shock. In both cases, we see an underreaction in expectations, which manifests itself in a positive forecast error after a shock that increases the forecasted variable, and vice-versa following a negative shock. After about 5-6 periods, the forecast error response, however, flips sign. This is exactly the eventual overreaction, mentioned above and documented in Angeletos et al. (2021). Thus, my model of inflation expectations matches these empirical findings.

Figure 12: Impulse Response Functions of Forecast Errors

Note: This figure shows the impulse-response functions of inflation forecast errors after a three-standard deviation positive (left) and negative (right) natural rate shock.

E.4 No Random Walk

A potential concern with the results stated in Section 3, in particular the *inflation-attention trap* in Figure 2, is that these findings are driven by the random walk assumption in the belief process of the agents. Relaxing the random-walk assumption requires to take a stand on the perceived average inflation. In this case, where I solve the model around the zero inflation steady state, this is quite innocuous. But later on, when I focus on Ramsey optimal policy, this cannot be done anymore without distorting the results, in the sense that agents might have a mean bias.

Figure 13 shows the same impulse response functions as reported in Figure 2 for the case of $\rho_\pi = 0.95$ and an average inflation of 0. We see a similar pattern, even though somewhat less pronounced. Inflation is persistently lower under limited attention due to slowly-adjusting inflation expectations. Expectations are updated even more sluggishly when $\rho_\pi < 1$. Further, this also dampens the initial response in inflation expectations, and thus, of inflation itself. Therefore, the attention trap is somewhat mitigated and the economy escapes the lower bound faster than with $\rho_\pi = 1$. Nevertheless, the nominal interest rate is low for longer due to the slow recovery of inflation.

Optimal policy with a bias in inflation expectations. In the main analysis, I have assumed that agents believe that inflation follows a random walk. Under this assumption, inflation expectations and inflation coincide on average. In the following, I relax this assumption and assume that the perceived persistence parameter is less than 1, $\rho_\pi < 1$. As discussed earlier, this yields the following inflation-expectations formation

$$\pi_{t+1|t}^e = (1 - \rho_\pi)\bar{\pi} + \rho_\pi \pi_{t|t-1}^e + \rho_\pi \gamma \left(\pi_t - \pi_{t|t-1}^e \right),$$

where $\bar{\pi}$ captures the long-run expectations of the agent. I set $\rho_\pi = 0.95$ and compare economies with different $\bar{\pi}$, namely $\bar{\pi} \in \{0\%, 2\%, 4\%\}$ (annualized).

Figure 14 shows the optimal inflation target (left panel) and welfare (12) (right panel) under Ramsey optimal policy for different levels of attention and different mean beliefs, $\bar{\pi}$. The blue-dashed lines show the results for the case with $\rho_\pi = 1$ (which is the baseline case discussed above), the gray-dashed-dotted lines show the results for $\rho_\pi = 0.95$ and $\bar{\pi} = 0\%$, the black-solid lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 2\%$, and the red-dotted lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 4\%$.

We see that introducing a mean bias in general leads to an increase in the optimal inflation target and additional welfare losses, independent of $\bar{\pi}$. This mainly comes from the fact that ρ_π is now below 1, which dampens the degree of updating captured by γ. Thus, once the economy gets stuck at the ELB and the policymaker tries to decrease real rates by increasing inflation expectations, actual inflation needs to increase more strongly. Therefore, a lower ρ_π

Figure 13: Impulse Response Functions to a Negative Natural Rate Shock
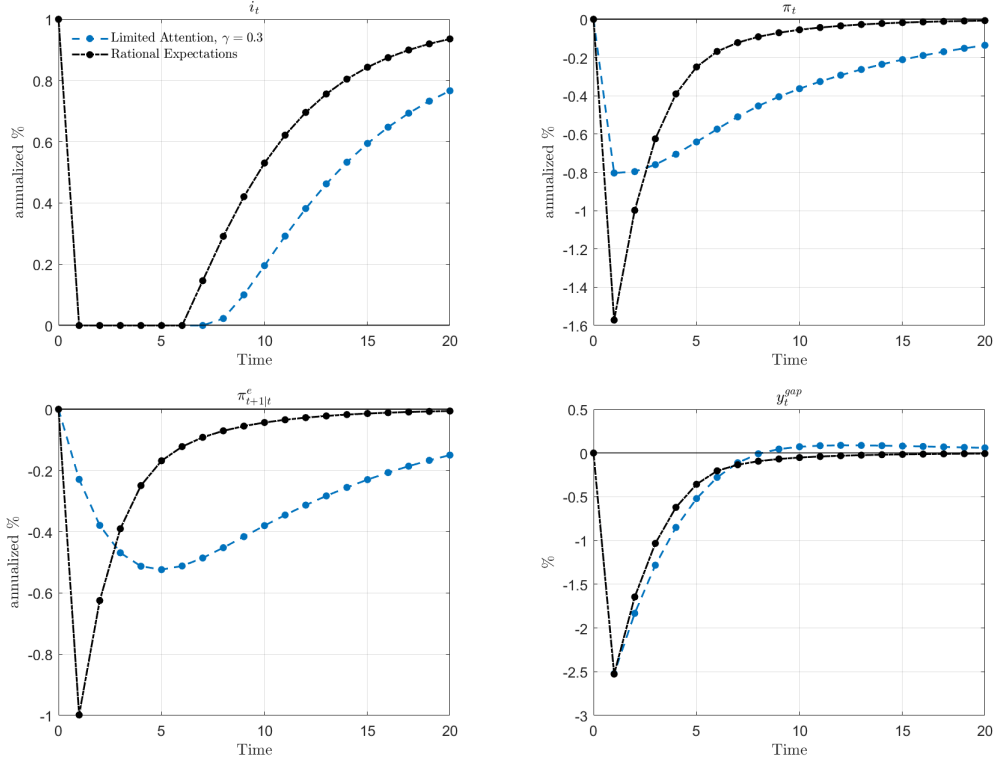
Note: This figure shows the impulse-response functions of the nominal interest rate (upper-right panel), inflation (upper-left panel), inflation expectations (lower-right) and the output gap (lower-left) to a negative natural rate shock of three standard deviations. The blue-dashed lines show the case for the limited-attention model and the black-dashed-dotted lines for the rational expectations model. Everything is in terms of percentage deviations from the respective steady state levels, expect the nominal rate is in levels.

can exacerbate attention traps when they occur.

Interestingly, the relationship between the optimal target and $\bar{\pi}$ is non-monotonic in the level of attention. While, for example, at $\gamma = 0.2$, the optimal target is highest at $\bar{\pi} = 0\%$, it is highest at $\bar{\pi} = 4\%$ when $\gamma = 0.05$. To understand this, we can write the unconditional average inflation expectations as

$$\pi^e = \frac{(1 - \rho_\pi)\bar{\pi} + \rho_\pi \gamma \pi}{1 - \rho_\pi(1 - \gamma)}.$$

The following Lemma sheds light on how $\bar{\pi}$ matters for average inflation expectations and how this depends on the level of attention, γ.

Lemma 11. *For the case $\rho_\pi = 1$, average inflation expectations move one-for-one with average inflation, independent of γ:*

$$\pi^e = \pi.$$

For the case $0 < \rho_\pi < 1$, average inflation expectations move less than one-for-one with average

Figure 14: Mean Bias, Optimal Inflation Target and Welfare
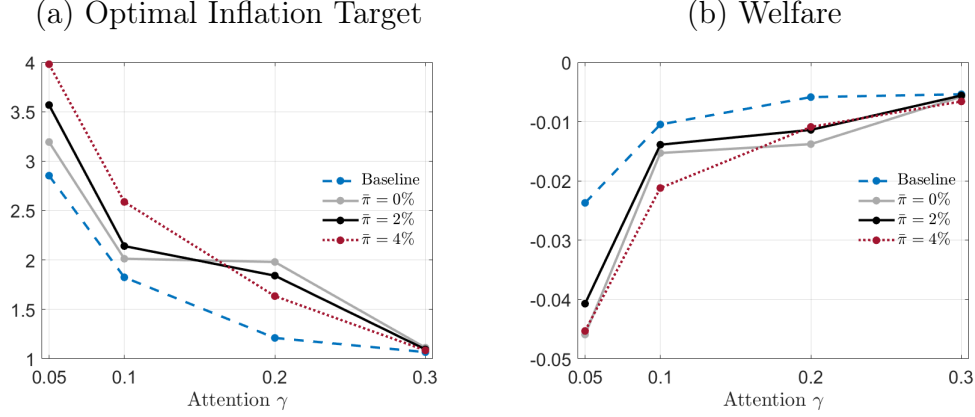
(a) Optimal Inflation Target

(b) Welfare

Notes: This figure shows the average inflation rate under Ramsey optimal policy (left panel) and welfare (12) (right panel) under Ramsey optimal policy for different levels of attention and different mean beliefs, $\bar{\pi}$. The blue-dashed lines show the results for the case with $\rho_\pi = 1$ (which is the baseline case), the gray-dashed-dotted lines show the results for $\rho_\pi = 0.95$ and $\bar{\pi} = 0\%$, the black-solid lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 2\%$, and the red-dotted lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 4\%$.

inflation

$$0 < \frac{\partial \pi^e}{\partial \pi} = \frac{\rho_\pi \gamma}{1 - \rho_\pi(1-\gamma)} < 1,$$

and the strength of this dependency increases with γ

$$\frac{\partial^2 \pi^e}{\partial \pi \partial \gamma} > 0.$$

Average inflation expectations move less than one-for-one with $\bar{\pi}$

$$0 < \frac{\partial \pi^e}{\partial \bar{\pi}} = \frac{(1 - \rho_\pi)}{1 - \rho_\pi(1-\gamma)} < 1,$$

*and the strength of this dependency decreases with γ*

$$\frac{\partial^2 \pi^e}{\partial \bar{\pi} \partial \gamma} < 0.$$

So, as attention falls, there are several opposing forces at work. On the one hand, the effect of $\bar{\pi}$ on average inflation expectations becomes stronger and thus, also exerts more pressure on actual inflation via the Phillips Curve. On the other hand, increasing the inflation target—average inflation—has a smaller effect on average inflation expectations at low levels of attention. Thus, to increase inflation expectations in this case, the inflation target needs to increase more strongly, which is of course costly. Comparing the optimal inflation targets in Figure 14, we see that at low levels of attention the first effect dominates. If $\bar{\pi}$ is relatively high, the inflation target is high.
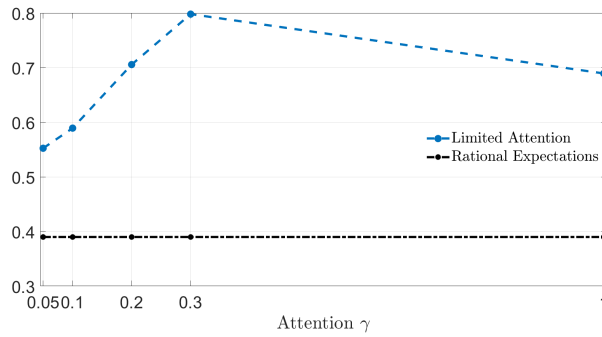
E.5 Figures to Section 4.3.5
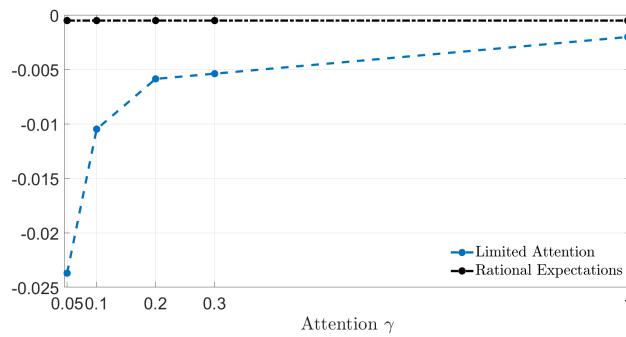
Figure 15: Full Attention, $\gamma = 1$

(a) Optimal Inflation Target
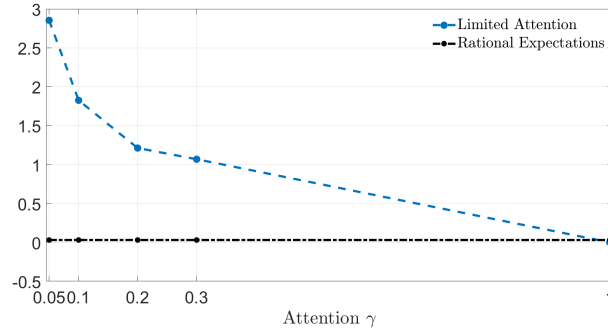
(b) Inflation Volatility

(c) Welfare

Notes: This figure shows the optimal inflation target (panel (a)), inflation volatility (panel (b)) and welfare (panel (c)) under Ramsey optimal policy for different levels of attention, including full attention, i.e., $\gamma = 1$ and compares it to the full-information rational expectations counterparts (black-dashed-dotted lines).