The Inflation Attention Threshold and Inflation Surges

Oliver Pfäuti*

First Version: August 2023

This version: November 2023

Link to most recent version

Abstract

At the outbreak of the recent inflation surge, the public's attention to inflation was low but increased rapidly once inflation started to rise. Using survey inflation expectations, I quantify when and by how much the public's attention to inflation changes. I estimate an attention threshold at an inflation rate of 4%, and that attention doubles when inflation exceeds this threshold. Negative supply shocks are more than twice as inflationary in the high-attention regime, and the increase in people's attention to inflation in early 2021 accounts for about 50% of the supply-driven inflation between 2021 and 2023. I develop a general equilibrium model accounting for the attention threshold and show that shocks that are usually short lived can lead to a persistent surge in inflation if they induce an increase in people's attention. The attention threshold leads to an asymmetry in the dynamics of inflation as periods of high inflation become substantially more likely than periods of low inflation.

JEL Codes: E3, E4, E5, E7

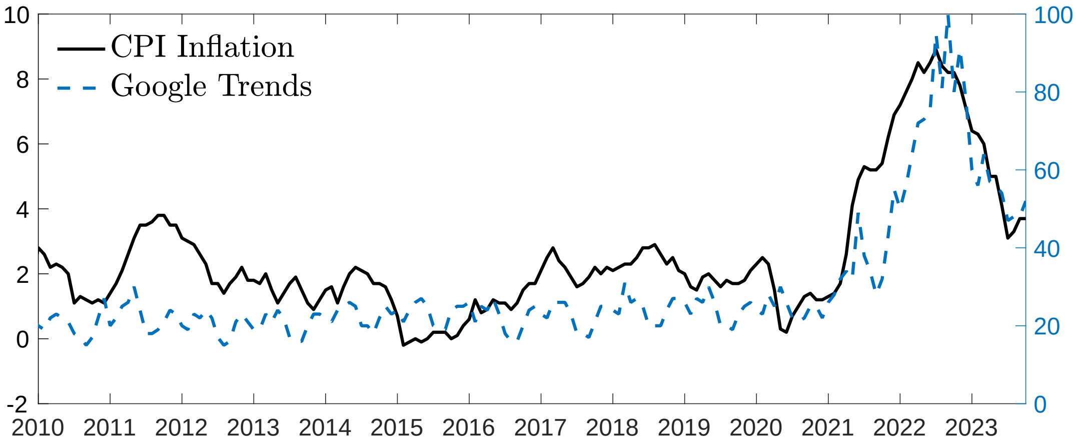
Keywords: Inattention, Inflation, Inflation Expectations, Monetary Policy

*University of Bern, Department of Economics, Schanzeneckstrasse 1, 3012 Bern, Switzerland, pfaeuti.oliver@gmail.com. I thank Luca Benati, Pierpaolo Benigno, Oli Coibion, Jose-Elias Gallegos, Laura Gáti, Niklas Kroner, Cyril Monnet, David Munro, Dirk Niepelt, Fabian Seyrich, Hannes Twieling and seminar participants at the University of Bern for helpful comments and suggestions.

1 Introduction

Inflation is back. After decades of low and stable inflation, inflation surged in many advanced economies during the recovery phase of the pandemic. Inflation turned out to be higher and more persistent than many expected.[1] With inflation rising, the public's attention to inflation increased as well (Figure 1). But is this increased attention to inflation merely a byproduct or a driver of high and persistent inflation?

Figure 1: Inflation is back... on people's minds



Notes: The black-solid line shows the monthly year-on-year US CPI inflation rate (left axis), and the blue-dashed line the number of Google searches in the US of inflation (right axis, normalized to have a maximum of 100).

To shed light on this question, I propose a way to estimate an inflation attention threshold at which attention changes, as well as the degrees of attention across the different attention regimes. Attention captures how strongly agents update their inflation expectations in response to forecast errors. I estimate the attention threshold and the attention levels in the two regimes jointly using survey data from the University of Michigan's Survey of Consumers for the period 1978 to 2023. I find an inflation attention threshold at an annualized inflation rate of 4.0%, and that attention is about 0.18 when inflation is below the threshold and doubles to 0.36 when inflation is above the threshold. An attention level of 0.36 means that following a 1pp forecast error, agents increase their inflation expectations by 0.36pp. Using the Survey of Consumer Expectations from the New York Fed for the period 2013-2023, I obtain similar results: attention in times when inflation was below 4% is about 0.21 and increases to 0.4 when inflation is above 4%. The data suggests that there is only one attention threshold, and once I control for this threshold, I do not find any evidence that attention increases with inflation *within* regime.

The degree of people's attention matters for the dynamics of inflation and inflation expectations. Using the high-frequency identification of Känzig (2021), I use oil supply news

[1] For example, Federal Reserve Chair Jerome H. Powell said in his 2021 Jackson Hole speech that inflation concerns are "likely to prove temporary" (see https://www.federalreserve.gov/newsevents/speech/powell20210827a.htm).

shocks as my proxy for supply shocks and I find that inflation increases more than twice as much in response to these shocks when they hit in the high-attention regime compared to the low-attention regime. A one standard deviation shock, that pushes inflation up by about 25 basis points on average across regimes, leads to a peak increase in inflation of more than 60 basis points when the shock hits in the high-attention regime. These effects are persistent as they only diminish after about one to two years. The effects on inflation expectations are even longer lived: as inflation starts to return to its initial value, inflation expectations start to overshoot actual inflation. This delayed overshooting happens later in the high-attention regime, indicating that inflation expectations stay persistently higher in the high-attention regime even after inflation already started to come down.

My findings imply that the US economy entered the high-attention regime most recently in early 2021. I show that supply shocks explain about 60% of the inflation surge from the time the US economy entered the high-attention regime in early 2021 until the peak in mid-2022. Without the attention increase, however, the implied inflation increase would have only been half as strong.[2] These results suggest that people's higher attention to inflation is not merely a byproduct but an important driver of the recent inflation surge and the increase in attention substantially amplified the already inflationary supply shocks.

To understand these findings as well as their implications for monetary policy, I develop a New Keynesian model accounting for the attention threshold and the changing levels of attention. I show analytically how the attention threshold alters the dynamics of inflation. Following a cost-push shock that pushes the economy into the high-attention regime, both the aggregate supply and aggregate demand curve become steeper—thus, the attention threshold gives rise to a *dynamic* non-linearity in the Phillips Curve and the aggregate demand curve. Additionally, the effects of the ongoing shock are amplified through the heightened attention level, leading to a further upward shift of the aggregate supply curve. On top of that, this shift occurs along a steeper aggregate demand curve, such that the inflationary effects become larger. Overall, this stylized example illustrates how an upward shift in attention may lead to self-reinforcing inflation dynamics without large changes in the output gap, and that inflation becomes more sensitive to shocks when attention is high, as found empirically.

I then calibrate the model to match the estimated attention levels in the two regimes and the attention threshold. Following an inflationary supply shock that pushes inflation above the attention threshold, inflation keeps on increasing through the endogenous attention increase. As attention increases, agents update their inflation expectations more strongly in response to the increase in inflation. These higher inflation expectations then fuel further inflation increases,

[2] I show that also other shocks, i.e., the *inflation shock* from Angeletos et al. (2020a) as well as monetary policy shocks, become more inflationary in the high-attention regime, indicating that shocks generally become more inflationary when attention to inflation is high. However, for the recent inflation surge I focus on oil supply news shocks as these shocks are available until the end of 2022.

leading to even higher inflation expectations. This illustrates how what usually would be a transitory inflation shock can become a very persistent one due to the heightened attention. As the shock dies out, inflation starts to decline after some time. However, inflation may remain elevated for a substantial period of time, because once it falls back below the threshold, people pay less attention to inflation, and hence, only slowly revise their inflation expectations downward. As their prior expectations are now relatively high due to the experienced high inflation period, expectations remain persistently high, leading to a slow decline of actual inflation. I show that these patterns of inflation and inflation expectations are consistent with the recent inflation surge in the US: inflation followed a hump-shaped pattern and inflation expectations initially fell short of the actual inflation surge, but then surpassed it once inflation started to decline.

That inflation expectations stay persistently high, even when inflation has already fallen back below the threshold, gives rise to a heightened risk of another subsequent inflation surge. The higher inflation expectations keep actual inflation higher for longer, and therefore, closer to the attention threshold. Thus, a subsequent inflationary shock is more likely to push inflation back above the threshold and therefore, leading to another episode of persistently high inflation.

The attention threshold induces an asymmetry in the dynamics of inflation: attention increases when inflation is particularly high but remains constant when inflation is particularly low. Relative to the model abstracting from the attention threshold, this asymmetry increases the risk of periods of persistently high inflation rates but leaves the risk of particularly low inflation rates unchanged.

This asymmetry also matters for the model's normative implications. When monetary policy sets the nominal interest rate following a simple and relatively dovish Taylor rule, the associated welfare losses are significantly larger compared to optimal policy rules or a strict inflation targeting rule.[3] The reason is that under a standard calibration of the Taylor rule, the economy spends a substantial amount of time in the high-attention regime in which inflation is high and volatile. Due to the asymmetry of the attention threshold the *average level of inflation* is higher when the economy is in the high-attention regime frequently. This differs from the model without the threshold or the one with fully-informed rational agents, where inflation fluctuates symmetrically around zero. This asymmetry of the attention threshold also offers a potential explanation for why we did not observe long lasting deflationary periods, for example, during the Great financial crisis (Coibion and Gorodnichenko, 2015). In fact, annualized quarter-on-quarter CPI inflation was negative only 7% of the time between 1978 and 2023, whereas it exceeded the attention threshold of 4.0% about 30% of the time during the same period.

[3]This particular finding is similar to what Gáti (2022) finds when studying the optimal monetary policy in a model with time-varying degrees of anchoring of long-run inflation expectations.

Related literature. This paper contributes to the ongoing debate about the drivers of the recent inflation surge and offers an alternative and novel perspective by highlighting the role of people's heightened attention during that period. Shapiro (2023) decomposes demand- and supply-driven inflation and estimates a similar role for supply factors as I do. Gagliardone and Gertler (2023) estimate a New Keynesian model to account for the recent inflation dynamics and attribute large parts of it to oil supply shocks and easy monetary policy. Bernanke and Blanchard (2023) find that shocks to prices given wages, including shocks to energy prices and sectoral shortages, played a key role for the recent inflation spike. Benigno and Eggertsson (2023) find evidence for a non-linear Phillips Curve that is steeper when labor markets are tight. Relatedly, Ball et al. (2022) find that the tightening of labor markets was a main driver of the rise in core inflation. Amiti et al. (2023) focus on the role of supply chain disruptions, labor supply constraints and the shift of consumption from services to goods, whereas Bianchi and Melosi (2022) and Bianchi et al. (2023) attribute the inflation surge to unfunded fiscal shocks or changes in people's beliefs about the fiscal framework. In contrast to all of these papers, I focus on the role of changes in people's attention to inflation and show that the increase in attention in early 2021 likely doubled the inflationary effects of supply shocks between early 2021 and the end of 2022.

Pfäuti (2023) documents that before the recent inflation surge, attention to inflation was at a historical low.[4] Bracha and Tang (2023) find that when inflation increases, attention to inflation increases as well. Korenok et al. (2022) find, for a large number of countries, that people's attention to inflation increases with inflation only after inflation exceeds a certain threshold. Using Google search data for the period from 2004 to 2022, they estimate this attention threshold to be at an annualized inflation rate of 3.55% for the US. Cavallo et al. (2017) show that survey respondents in high-inflation environments (Argentina) respond less to information about inflation than in low-inflation environments (United States), which is consistent with higher attention to inflation in high-inflation environments. Weber et al. (2023) confirm, using a range of randomized control trials spanning over several years and different countries, that attention of households and firms is indeed higher in times of high inflation. Kroner (2023) focuses on financial markets and shows that attention—measured as asset price responses to inflation news—is higher in times of high inflation. My key innovation relative to

[4] For the years before the recent inflation surge, Candia et al. (2021, 2023) and Coibion et al. (2020) show that U.S. firms as well as households are usually poorly informed about and inattentive to inflation and monetary policy (see also Weber et al. (2022) for a recent review). Goldstein (2023) also finds that inattention varies over time. Different forms of changing attention are considered, e.g., in Kim and Binder (2023) who examine how repeat participation in surveys that ask about inflation expectations may lead to higher attention to inflation, or in Flynn and Sastry (2022) who show—by using a textual proxy for firms' attention toward macroeconomic conditions—that attention is counter-cyclical, or Gallegos (2023) who shows that firms' less sluggish inflation expectations after the Great inflation period offer a potential explanation for the decrease in the persistence of inflation as found, e.g., in Benati (2008).

these papers is that I provide estimates of the attention threshold and the attention levels in the two regimes in a way that directly maps into otherwise standard macroeconomic models, and I quantify the role of these attention changes in the recent inflation surge.

I further contribute to the literature on the state dependency of shocks (Auerbach and Gorodnichenko, 2012a,b; Caggiano et al., 2014; Ramey and Zubairy, 2018; Tenreyro and Thwaites, 2016). The two papers most closely related are Ascari and Haber (2022) and Joussier et al. (2023). Ascari and Haber (2022) show that the inflationary effects of monetary policy shocks are larger in times of high trend inflation. Joussier et al. (2023) find that French firms pass through energy price shocks more to their prices in times of high inflation.[5] I contribute to that literature by showing that attention to inflation matters for the transmission of supply and monetary shocks to aggregate inflation and inflation expectations.

The theoretical insights in this paper contribute to a growing literature on the role of changes in attention and the degree of anchoring of inflation expectations. Evans and Ramey (1995) propose a model in which agents choose how forward-looking they want to be and show that inflation remains stable when agents are not very forward looking.[6] Gáti (2022) studies how changes in the degree of anchoring of long-run inflation expectations affect the optimal monetary policy. In her paper, anchoring changes continuously, thus, can also jump in response to large shocks. In contrast, attention in my model only changes across but not within regime, which is consistent with what I find in the data, namely, that attention is constant within regimes. Carvalho et al. (2022) focus on discrete changes in anchoring of long-run inflation expectations, but do so in a partial equilibrium setting. In Pfäuti (2023), I study the implications of low attention for optimal monetary policy when the zero lower bound poses a constraint to monetary policy. In contrast to that paper, I allow here for an inflation attention threshold, whereas Pfäuti (2023) compares economies with different—but time-invariant—degrees of attention. Hazell et al. (2022) show that the greater inflation stability after 1990 is mostly due to more firmly-anchored long-run inflation expectations. Related, Afrouzi and Yang (2020) show that in a model of dynamic rational inattention the stance of monetary policy affects the slope of the Phillips curve (see also Jørgensen and Lansing (2023)). The main contribution of the present paper to this literature is that I incorporate and estimate an inflation attention threshold into a general equilibrium model and show how accounting for such a threshold can help us better understand how inflation surges may happen and how they can sustain themselves, as well as

[5] Joussier et al. (2023) use Google Trends data to proxy for attention to energy prices and energy cost and show that the pass through of energy shocks to prices in France is stronger in times of higher attention to energy prices.

[6] Focusing on endogenous but constant attention, Mackowiak and Wiederholt (2009) show how firms set prices optimally information is costly, and Paciello and Wiederholt (2014) how the optimal monetary policy changes when firm managers decide optimally how much attention they want to pay to aggregate conditions compared to a setting where their attention is exogenous (see also Sims (2003, 2010)). Reis (2006a,b) study how decisions of consumers and producers change when updating their information set infrequently.

the monetary policy implications that follow from such attention-fueled inflation surges.

Outline. In Section 2, I estimate the inflation attention threshold and attention across regimes. In Section 3, I show that the attention regime matters for the inflationary effects of shocks, and quantify the role of attention in the recent inflation surge. I introduce the New Keynesian model with limited attention in Section 4, and the model's positive results in Section 5. In Section 6, I derive the model's normative implications and Section 7 concludes.

2 Attention and the Inflation Attention Threshold

In this section, I show how I estimate the inflation attention threshold and the different degrees of attention in the two attention regimes, and present the estimation results. To arrive at these results, I extend the method I develop in Pfäuti (2023) by allowing for an inflation-attention threshold.[7] In Appendix A, I show how a simple model of optimal attention choice provides a microfoundation for the following results.

The agent believes that inflation π follows

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

where $\underline{\pi}$ is the agent's long-run belief about inflation and ρ_π is the perceived persistence of inflation.[8] I assume that the error term ν_t is normally distributed with mean zero and variance σ_ν^2. The agent receives a signal about inflation of the form

$$s_t = \pi_t + \varepsilon_t,$$

where the noise ε_t is assumed to be normally distributed with variance σ_ε^2. Higher attention is reflected in less noise, i.e., a lower σ_ε^2.

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

$$\tilde{E}_t \pi_{t+1} = (1 - \rho_\pi)\underline{\pi} + \rho_\pi \tilde{E}_{t-1}\pi_t + \rho_\pi \gamma_\pi \left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, \tag{1}$$

where the updating gain γ_π captures how attentive the agent is and crucially depends on the noise variance σ_ε^2 (see Appendix A for details). When the agent is relatively inattentive, she

[7] The method proposed in Pfäuti (2023) builds on Mackowiak et al. (2023) and Vellekoop and Wiederholt (2019).

[8] Agents' subjective model need not necessarily be consistent with the actual law of motion (see Andre et al. (2022) and Macaulay (2022) for empirical evidence that the subjective models agents hold may not necessarily be consistent with the actual behavior of the economy or with experts' subjective models).

obtains relatively noisy signals and thus, puts little weight on these signals, reflected in a small γ_π. Therefore, lower attention, i.e., a lower γ_π, implies that the agent updates her expectations to a given forecast error, $(\pi_t - \tilde{E}_{t-1}\pi_t)$, less strongly.

In principle, attention to inflation may vary and depend on the economic environment.[9] To account for such changes in people's attention, I allow attention to change when inflation exceeds a certain threshold $\bar{\pi}$:

$$\tilde{E}_t\pi_{t+1} = \begin{cases} (1-\rho_\pi)\underline{\pi} + \rho_\pi \tilde{E}_{t-1}\pi_t + \rho_\pi \gamma_{\pi,L}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, \text{ when } \pi_{t-1} < \bar{\pi} \\ (1-\rho_\pi)\underline{\pi} + \rho_\pi \tilde{E}_{t-1}\pi_t + \rho_\pi \gamma_{\pi,H}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, \text{ when } \pi_{t-1} > \bar{\pi}. \end{cases} \tag{2}$$

Later on, I test for multiple thresholds and changes of attention within regime.

One potential reason for such changes in attention may arise from changes in how costly it is to acquire and process information about inflation. Consistent with that, I show in Figure 14 in Appendix B that news coverage of inflation is substantially higher in times of higher inflation. Thus, information about inflation is more prevalent in times of higher inflation, which lowers the cost of acquiring information about inflation, leading to an increase in people's attention.[10] Hence, when there is more news about inflation in regime H, attention is higher: $\gamma_{\pi,H} > \gamma_{\pi,L}$. In the estimation later on, I will test this prediction of the theory rather than imposing it. I assume that agents do not foresee these changes in news reporting and form their expectations within regime r as if the cost of information would remain unchanged.[11]

2.1 Estimating attention and the attention threshold

Data. As my measure of inflation expectations, I rely on the Survey of Consumers from the University of Michigan. For my baseline specification, I use average and median household inflation expectations for the period 1978-2023. For the period 2013-2023, I also use individual inflation expectations from the Survey of Consumer Expectations from the New York Fed. Even though I focus on inflation and inflation expectations over one quarter, I use monthly

[9]The microfoundation in Appendix A predicts that attention depends positively on the agent's perceived persistence and volatility of inflation, preferences, and negatively on how costly it is to acquire and process information about inflation.

[10]Consistent with this, Bracha and Tang (2023) show that higher inflation rates indeed lead to more media reporting about inflation, and Lamla and Lein (2014) find that more intensive news reporting about inflation improves the accuracy of consumers' inflation expectations, consistent with agents being more attentive. Schmidt et al. (2023) show that during episodes of intensive newspaper coverage of inflation, news reporting has strong effects on inflation expectations but not during other episodes. Larsen et al. (2021) find that news media coverage predicts households' inflation expectations, and Nimark and Pitschner (2019) show that major events (such as strong inflation increases) lead to a shift in the news focus towards these events.

[11]Furthermore, I assume that the agents always use the steady state Kalman filter within regime to form their expectations, which is a standard assumption in the rational inattention literature and basically means that the agent receives all her signals before forming her expectations (see, e.g., Mackowiak and Wiederholt (2009); Maćkowiak et al. (2018, 2020)).

data to increase the number of observations. As a robustness check, however, I also consider expectations at quarterly frequency. One advantage of using quarterly observations is that the Survey of Consumers provides mean expectations going back to 1960Q2. Both surveys ask consumers for their price growth expectations one-year ahead: $\tilde{E}_t \pi_{t+12}$. I compute one-quarter-ahead forecasts by dividing the one-year-ahead forecasts by four: $\tilde{E}_t \pi_{t+3} \equiv \frac{\tilde{E}_t \pi_{t+12}}{4}$. Using $\tilde{E}_t \pi_{t+3} \equiv \left[\left(1 + \tilde{E}_t \pi_{t+12} \right)^{\frac{1}{4}} - 1 \right]$ instead does not change the results. Computing one-quarter-ahead expectations allows me to compare the results directly to the model which is calibrated at quarterly frequency. For actual inflation, I use the monthly CPI inflation rate from the FRED database and to be consistent with the model, I focus on quarter-on-quarter inflation: $\pi_t \equiv \frac{P_t - P_{t-3}}{P_{t-3}}$. Figure 13 in Appendix B plots the time series of these variables, and in Appendix B I also discuss heterogeneity in attention across demographic groups and show that the following results are similar when focusing on the Covid period.

Estimating attention and the attention threshold. In order to estimate the two attention levels $\gamma_{\pi,r}$ for $r \in \{L, H\}$, as well as the attention threshold $\bar{\pi}$, I estimate the following threshold regression:

$$\begin{aligned} \tilde{E}_t \pi_{t+3} = & \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}} \left(\beta_{0,L} + \beta_{1,L} \tilde{E}_{t-3} \pi_t + \beta_{2,L} \left(\pi_t - \tilde{E}_{t-3} \pi_t \right) \right) \\ & + \left(1 - \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}}\right) \left(\beta_{0,H} + \beta_{1,H} \tilde{E}_{t-3} \pi_t + \beta_{2,H} \left(\pi_t - \tilde{E}_{t-3} \pi_t \right) \right) + \epsilon_t, \end{aligned} \tag{3}$$

where $\beta_{0,r} = (1 - \rho_{\pi,r}) \underline{\pi}_r$ denotes the intercept in regime $r \in \{L, H\}$, $\beta_{1,r} = \rho_{\pi,r}$ the perceived persistence, and $\frac{\beta_{2,r}}{\beta_{1,r}} = \gamma_{\pi,r}$ the attention level in regime r. $\mathbb{1}_{\pi_{t-1} \leq \bar{\pi}}$ is the indicator function that equals one when inflation in the previous month π_{t-1} was below the threshold $\bar{\pi}$ and zero otherwise. The threshold value is then estimated jointly with the regression coefficients by minimizing the sum of squared residuals obtained for all possible thresholds (see, e.g., Gonzalo and Pitarakis (2002) and Hansen (2011)). Note, that I do not impose that the attention level in the regime in which inflation is above the threshold needs to be higher than in the regime with inflation below the threshold.

Estimation results. Table 1 shows the estimation results. For the baseline specification, I estimate an attention threshold of $\hat{\bar{\pi}} = 3.98\%$. Attention in the regime in which inflation is below this threshold is $\widehat{\gamma}_{\pi,L} = 0.18$, and attention in the high-inflation regime is equal to $\widehat{\gamma}_{\pi,H} = 0.36$. Thus, attention in the high-inflation regime is higher than in the low-inflation regime, and I therefore refer to the regime with inflation above the threshold as the *high-attention regime*. This is consistent with the theory that higher attention may be driven by more extensive news coverage, since news coverage is substantially higher when inflation is above 4% (see Appendix B). The threshold value of 3.98% implies that the US economy spent about 31% of the time in the period 1978-2023 in the high-attention regime. Most recently, inflation exceeded this threshold in early 2021. Google searches also started to increase around that time (see Figure 1 in the Introduction). Using the Bayesian Information Criterion or the Hannan–Quinn information criterion to select the numbers of thresholds, the data prefers the specification with only one threshold.

The results for median expectations are similar. Attention in both regimes tends to be somewhat lower when using median expectations and the attention threshold higher. In both cases, the null hypothesis that the two attention levels across regimes are equal is rejected at all conventional significance levels (p-values of 0.000 in both cases, see last column).

Table 1: Estimated attention levels and the attention threshold

	Threshold $\hat{\bar{\pi}}$	Low Att. $\widehat{\gamma}_{\pi,L}$	High Att. $\widehat{\gamma}_{\pi,H}$	p-val. $H_0 : \gamma_{\pi,L} = \gamma_{\pi,H}$
Mean expectations	3.98%	0.18	0.36	0.000
s.e.		(0.013)	(0.037)	
Median expextations	4.41%	0.16	0.23	0.000
s.e.		(0.013)	(0.028)	
Quarterly frequency	3.21%	0.14	0.38	0.000
s.e.		(0.033)	(0.076)	

Notes: This table shows the results from regression (3), where $\hat{\bar{\pi}}$ denotes the estimated threshold, $\widehat{\gamma}_{\pi,L}$ and $\widehat{\gamma}_{\pi,H}$ the estimated attention levels when inflation is below or above the threshold, respectively. The last column shows the p-value for the null hypothesis that the two attention levels are equal. The upper half of the table presents the results when using average expectations, and the lower half when using median expectations. Standard errors are robust with respect to heteroskedasticity.

The last two rows in Table 1 show the results when using observations at quarterly frequency for the period 1960Q2-2023Q2. We see that the estimated threshold is somewhat lower at 3.21%. The estimated attention levels within regime are very similar to my baseline specification and again, the difference in attention across regimes is highly statistically significant.

When I use the current inflation rate rather than the lagged inflation rate as the threshold-defining variable, I find practically identical results. Attention is 0.18 when inflation is below 4% and increases to 0.36 when inflation is above 4%. When I use the average of the last three months as the threshold-defining variable, I estimate a threshold at 3.5% and attention levels of $\widehat{\gamma}_{\pi,H} = 0.19$ and $\widehat{\gamma}_{\pi,H} = 0.34$.

When using individual consumer inflation expectations from the Survey of Consumer Expectations, I obtain similar results even though this survey is only available since 2013. When last month's inflation was below the threshold of 4%, I estimate an attention level of 0.21. When inflation was above the threshold, attention doubles to 0.4, and I reject the nullhypothesis that the two estimates are equal (p-value of 0.000).

My estimated attention threshold is slightly higher than the one estimated in Korenok et al. (2022). They estimate for the US that once inflation exceeds the threshold of 3.55%

that attention—measured using Google searches of inflation—increases with inflation whereas it does not below the threshold. One reason why I find a higher threshold could be that people might start to google for information about inflation at lower levels of inflation already, but only start incorporating that information in their expectations once inflation further increases. Additionally, I focus on the time starting in 1978, whereas their sample is restricted to 2004-2022 because the Google search data is not available for years before 2004. They also consider news coverage of inflation, and in that case, they estimate a threshold between 3.77-3.94%, which is very close to the threshold I estimate. Overall my results align with their findings, and my approach provides a quantification of the threshold and the attention levels in a way that is directly applicable to standard macroeconomic models, as I show later.

Attention changes within regime. In Table 1, we saw that attention increases strongly when inflation exceeds the threshold of 4%. But what about changes within regime? To look at this, I estimate a time series of attention. In particular, I estimate the regression

$$\tilde{E}_t \pi_{t+3} = \beta_0 + \beta_1 \tilde{E}_{t-3}\pi_t + \beta_2 \left(\pi_t - \tilde{E}_{t-3}\pi_t\right) + \epsilon_t,$$

using a rolling-windows approach, where each window has a length of one year. I denote the estimated time series of attention parameters by $\widehat{\gamma}_{\pi,t}$, and I compute the window-specific average of the monthly quarter-on-quarter inflation rate, π_t^{avg}. To then test whether attention within regime is higher when inflation is higher, I estimate the following regression

$$\widehat{\gamma}_{\pi,t} = \delta_0 + \delta_1 \mathbb{1}_{\pi_t^{avg} \geq 4} + \delta_2 \pi_t^{avg} + \delta_3 \mathbb{1}_{\pi_t^{avg} \geq 4} \pi_t^{avg} + \varepsilon_t, \tag{4}$$

where $\mathbb{1}_{\pi_t^{avg} \geq 4}$ is an indicator that equals one when in period t average inflation over the last twelve months is above 4% and zero otherwise. Thus, δ_1 tells us the difference of attention across regimes, δ_2 the effect of inflation on attention and δ_3 the additional effect of inflation on attention in the high-inflation regime.[12]

The first row in Table 2 shows the results. We see that attention is significantly higher in the higher-inflation regime, as indicated by the estimate of δ_1. Yet, inflation does not have any additional significant effect on attention when accounting for the threshold, as depicted by the last two columns.

As a robustness check, I also estimate

$$\widehat{\gamma}_{\pi,t} = \delta_0 + \delta_1 \mathbb{1}_{\pi_t^{avg} \geq 4} + \delta_2 \pi_{t-1} + \delta_3 \mathbb{1}_{\pi_{t-1} \geq 4} \pi_{t-1} + \varepsilon_t, \tag{5}$$

where I use lagged inflation as the independent variable rather than average inflation. The

[12] To be consistent with the theory, I impose that attention is between 0 and 1, see expression (33).

lower part of Table 2 shows that the results are robust to this change. Once I control for the inflation threshold, I do not find any evidence for a positive relationship between inflation and attention on top of the attention shift across regimes.

Table 2: Attention changes within regime

	$\widehat{\delta}_1$	$\widehat{\delta}_2$	$\widehat{\delta}_3$
Regression (4)	0.393**	0.053	-0.079
s.e.	(0.192)	(0.047)	(0.051)
Regression (5)	0.119*	-0.010	0.010
s.e.	(0.0641)	(0.0141)	(0.0141)

Notes: This table shows the results from regression (4) (upper half of the table) and from regression (5) (lower half). Standard errors are robust with respect to heteroskedasticity and serial correlaton (Newey and West (1987) with 12 lags). Significance levels: *: p-value < 0.1, **: p-value < 0.05, ***: p-value < 0.01.

3 The Role of Attention for the Propagation of Shocks

In this section, I show that negative supply shocks become more inflationary when attention is high, and that the increase in attention in early 2021 substantially amplified the inflationary effects of supply shocks in the following inflation surge.

Supply shocks. As my empirical measure of supply shocks, I use the oil supply news shocks from Känzig (2021) for the period 1975M1-2022M12. In a first step, oil surprises are identified by using institutional features of the Organization of the Petroleum Exporting Countries (OPEC) and high-frequency data on variation in oil futures prices around OPEC announcements. In a second step, the resulting surprises are then used as an external instrument in an oil VAR, to identify a structural oil supply news shock. In the following, I show the responses to a negative shock that pushes up oil prices and lowers oil production. I refer to these shocks as oil news shocks, cost-push shocks or supply shocks interchangeably.

Later on, I show that the following results also hold when focusing on other shocks. Given that the oil supply news shocks cover most of the recent inflation surge—while the other shocks do not—I use the oil supply news shocks as my baseline.

Attention regimes. Given the results in Section 2, I define the high-attention regime to be periods in which inflation in the past month exceeded 4%. Later, I use Google Trends as an alternative indicator for the attention regime, with the drawback that this data is only available since 2004.

Table 3 provides some summary statistics of the supply shocks: the mean, the standard deviation and the number of observations N of the shocks for the high-attention and low-

Table 3: Shock properties across regimes

Regime	Mean	Standard deviation	N
Inflation as regime-defining variable			
High	-0.017	0.554	192
Low	0.001	0.550	384
Google Trends as regime-defining variable			
High	-0.026	0.6762	40
Low	0.005	0.629	187

Notes: This table shows the mean, the standard deviation and the number of observations for the oil supply news shocks across the two attention regimes. The upper part of the table shows the case where the previous month's inflation is the threshold-defining variable and the lower part the case where Google Trends are used to identify the two regimes.

attention regime, separately. The upper part of the table shows these statistics for the case in which the regimes are defined based on whether the previous month's inflation rate was below or above the 4% threshold, and the lower part of the table for the case where the regimes are defined based on Google Trends data. We see that in both cases, the shock series have similar first and second moments across regimes. Thus, the following results are unlikely to be driven by differences in the shock series across regimes.

The high-attention periods were especially the 1970s and 1980s as well as the recent inflation surge after the Covid-19 pandemic.[13] In the following, I will show that the implications of being in the high-attention regime for the transmission of supply shocks to inflation are very similar for these two periods.

Quantifying the role of attention for the propagation of supply shocks to inflation. To examine whether the effects of negative supply shocks on inflation differ across regimes, I estimate the local projection (Jordà, 2005):

$$y_{t+j} - y_{t-1} = \mathbb{1}_H \left(\alpha_j^H + \beta_j^H \varepsilon_t\right) + (1 - \mathbb{1}_H) \left(\alpha_j^L + \beta_j^L \varepsilon_t\right) + \Gamma' X_t + u_{t+j}, \tag{6}$$

where y_{t+j} denotes inflation (or inflation expectations) at time $t+j$, $\mathbb{1}_H$ is an indicator function that equals one when the economy is in the high-attention regime and 0 else, ε_t denotes the shock at time t and X_t are controls. In my baseline estimation, I use four lags of the shock, four lags of the unemployment rate and lags 2 to 4 of inflation as controls.[14] In Appendix C, I show that the results remain robust when using other controls.

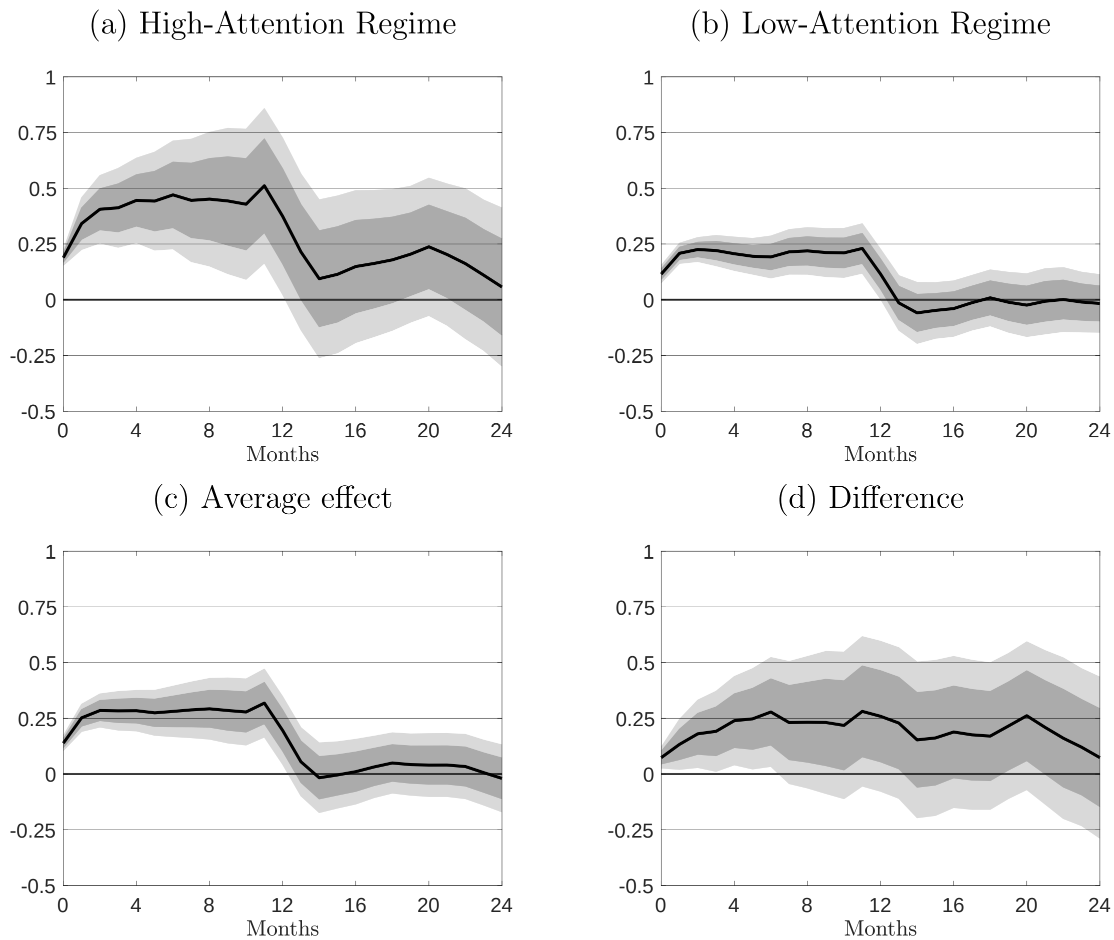
Figure 2 shows the estimation results of (6). Panel (a) depicts the inflation response to

[13]Consistent with my finding that the U.S. was in the high-attention regime during the 1970s and 1980s, Reis (2022) documents an unanchoring of inflation expectations in the Great Inflation period.

[14]I do not include the first lag of inflation as control because it already is on the left-hand side of regression (6). Including it as a control, however, barely affects the results.

a negative oil news shock in the high-attention regime, panel (b) in the low-attention regime, panel (c) shows the average effect, and panel (d) the difference between the effects in the high-attention regime and the low-attention regime. In all cases, I consider a one-standard deviation shock. The dark shaded areas depict the 68% confidence bands and the light-shaded areas the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Figure 2: Inflation response to an oil supply news shock

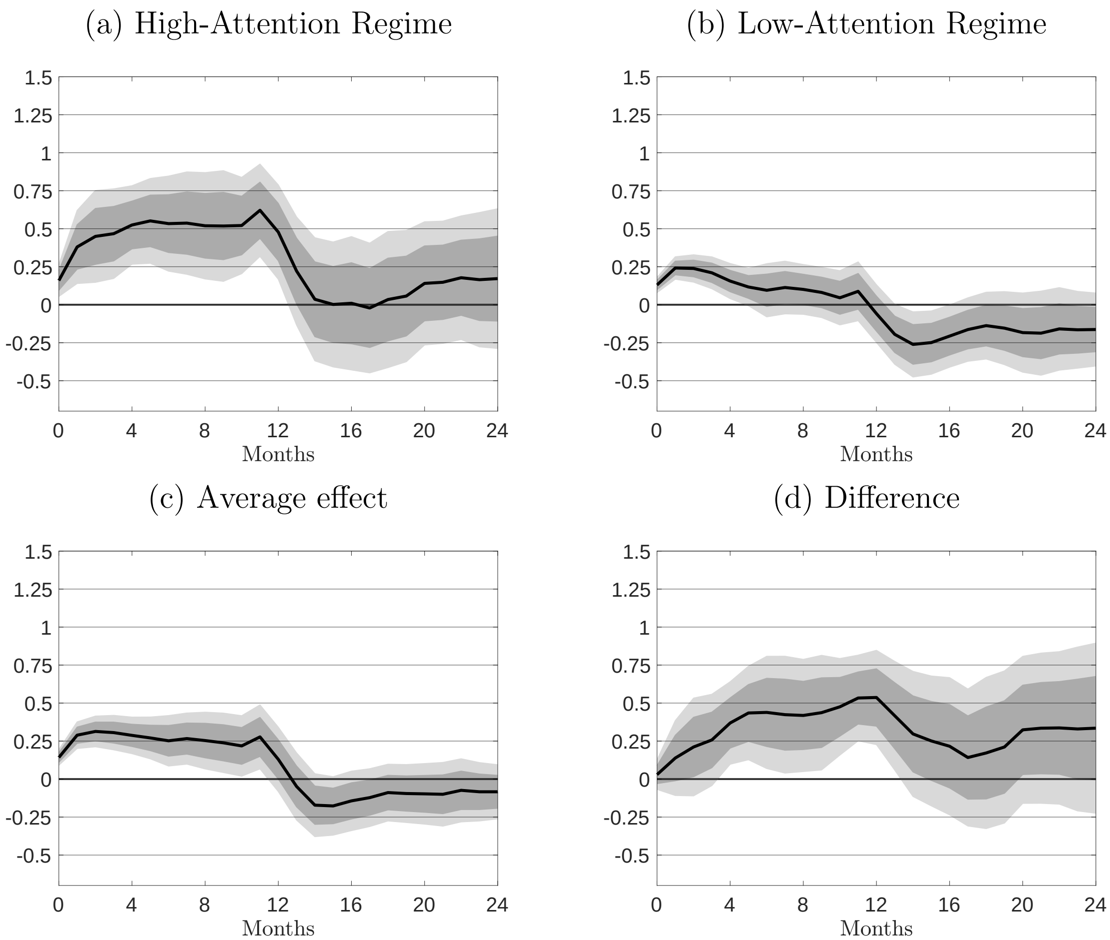


Notes: This figure shows the inflation response to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Inflation increases about twice as much in the high-attention regime compared to the low-attention regime. These differences are quite persistent and statistically significant at the 10% significance level in the first six months. Thus, when the economy is hit by a one-standard deviation supply shock when people's attention to inflation is high, inflation increases on average by about 50 basis points. In contrast, when attention to inflation is low at the time of the shock, inflation only increases by about 25 basis points.

Using Google Trends as the threshold-defining variable. Instead of using the lagged inflation rate as the threshold-defining variable, I now use Google Trends data to define the regimes. This data, however, is only available since 2004. Thus, the high-attention regime largely coincides with the recent inflation surge. The data is normalized such that the month with the most Google searches of the word "inflation" is assigned a value of 100 and all the other months are expressed relative to that month. I assign months to the high-attention regime when Google searches of inflation in that month exceed the 75^{th} percentile. I then estimate the local projection (6) for the period since 2004M1. I use the same control variables as before and I further include four lags of the number of Google searches as control variables.

Figure 3: Inflation response to an oil supply shock with Google data as threshold-defining variable



Notes: This figure shows the inflation response to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The attention regimes are defined based on the number of Google searches of the word inflation in the current month. The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Figure 3 shows the results. We see that they confirm the previous results that inflation responds substantially and significantly more to supply shocks when attention to inflation is high. In fact, the differences across regimes are even more pronounced than in Figure 2. While inflation again peaks at about 25 basis points in the low-attention regime, it peaks at more than 60 basis points in the high-attention regime (it peaks at about 50bp when considering the whole sample in Figure 2). These effects are also more persistent: one year after the shock, the inflation response in the high-attention regime is still more than 50 basis points higher than in the low-attention regime and this difference is highly statistically significant.

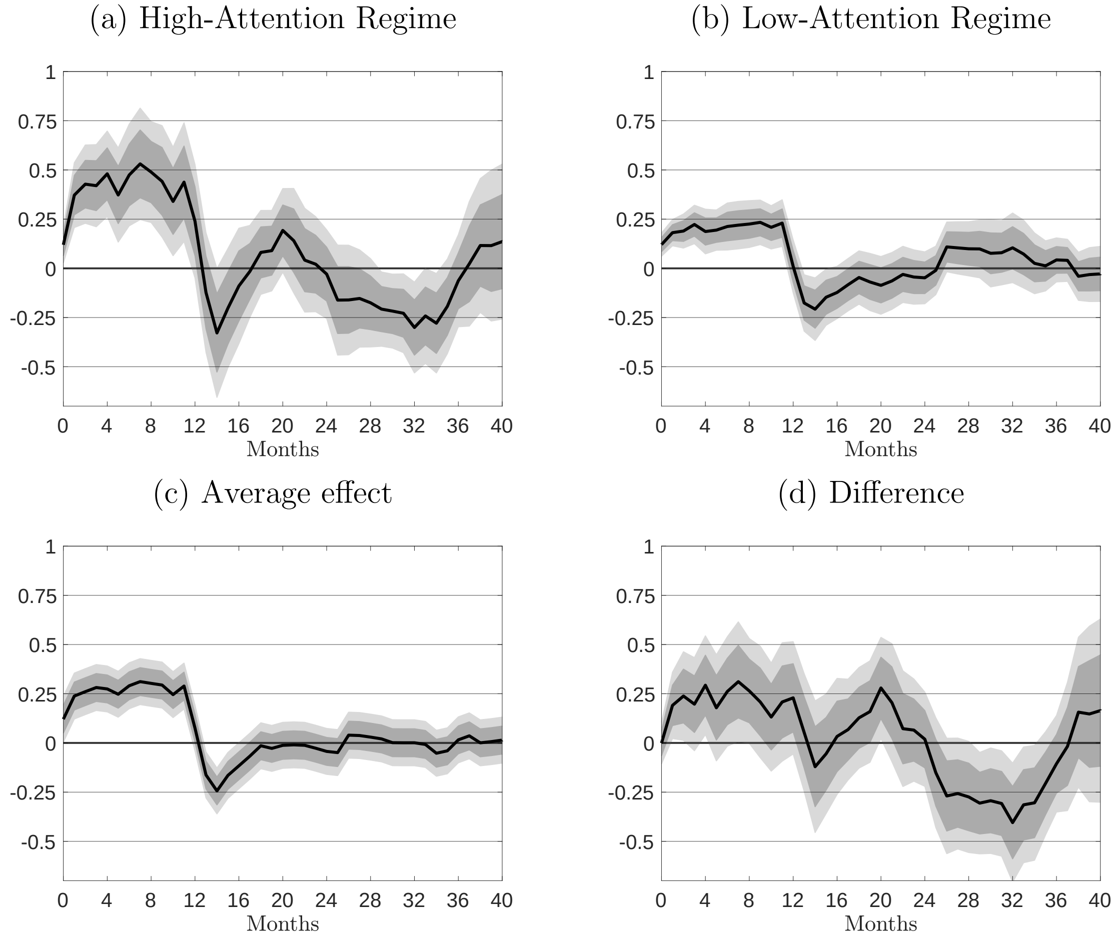
Forecast errors. Figure 4 shows the dynamics of forecast errors in response to an oil supply news shock, using average inflation expectations from the Survey of Consumers. In both regimes, forecast errors are initially positive—indicating an underreaction in expectations as a positive error means that inflation increased more than what was expected—followed by a delayed overshooting. This delayed overshooting is captured by the negative forecast errors at longer horizons. The key difference across regimes is that this delayed overshooting happens later in the high-attention regime, and significantly so, as panel (d) shows. This suggests that inflation expectations stay persistently higher in the high attention regime even after inflation already started to come down. When discussing the model results later, I will discuss how these persistently high inflation expectations give rise to a heightened risk of a subsequent high inflation period after an inflation surge.[15]

Other shocks. So far, I focused on the implications of supply shocks on inflation. In Appendix C, I show that other shocks become more inflationary in the high-attention regime as well. In figure 16, I show that inflation responds twice as much to the *main inflation shock* from Angeletos et al. (2020a) and that these differences are highly persistent. In figure 17, I show that inflation also responds more strongly in the high-attention regime to monetary policy shocks identified using a high-frequency identification (the shocks are taken from Jarociński and Karadi (2020) and are purged from the information effects of monetary policy statements). A drawback of using these monetary policy shocks, however, is that they are only available for the period 1990-2019. Thus, the sample does not include most of the high-inflation periods. Therefore, while the differences across regimes are substantial in magnitude, the differences are less statistically significant.

Robustness. In Appendix C, I discuss a number of robustness checks. In particular, I show that my results are not driven by different oil price responses (figures 18 and 19) or a higher sensitivity of macroeconomic variables in general (figures 20 and 21). I also show that the

[15] Angeletos et al. (2020b) also find that inflation and unemployment expectations show patterns of initial underreaction (mirrored by positive forecast errors) and delayed overshooting (negative forecast errors). They do not, however, distinguish between different attention regimes.

Figure 4: Forecast error response to an oil supply news shock



Notes: This figure shows the response of forecast errors (using mean household inflation expectations) to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The attention regimes are defined based on the number of Google searches of the word inflation in the current month. The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

results are very similar when excluding the Covid period (figure 22), when excluding the Great Inflation period (figure 23), when using different control variables (figure 24), or when focusing on the cumulative changes in the price level rather than the cumulative changes in inflation compared to the initial inflation rate (figure 25). Overall, the results in this section show that inflation responds 2-3 times as strongly to inflationary supply shocks when the public's attention to inflation is high compared to periods in which it is low.

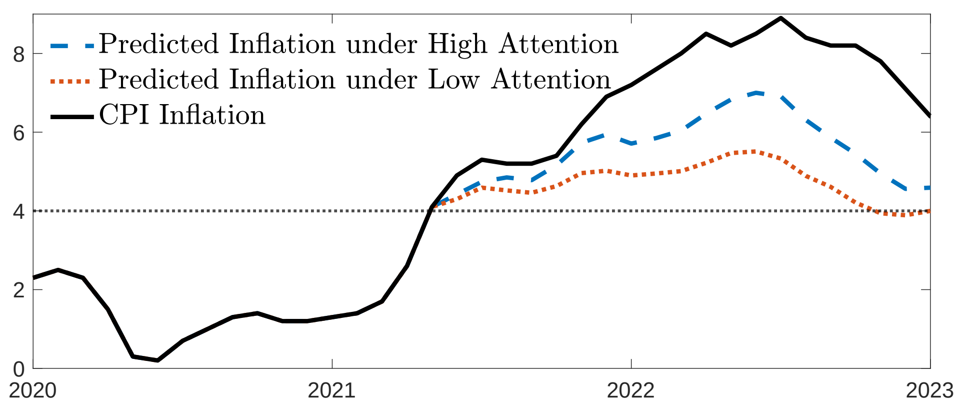
3.1 The Role of Attention in the Post-Covid Inflation Surge

I now show that oil supply news shocks as well as the increase in people's attention to inflation played a quantitatively important role in driving the post-Covid inflation surge in the United

States.[16] The black-solid line in Figure 5 shows the evolution of CPI inflation for the period 2020-2023.

To quantify the importance of the change in attention and of oil supply shocks, I estimate the counterfactual evolution of CPI inflation when only considering oil supply news shocks starting in April 2021, which is when inflation exceeded 4% and thus, the economy entered the high-attention regime.[17] I do this twice: once, taking the increase in attention into account, and once, without the increase in attention. For the case with the increase in attention, I therefore feed in the identified oil supply shocks from April 2021 until the end of 2022 using the estimated impulse response functions from panel (a) of Figure 2. The blue-dashed line shows the implied inflation rate. We see that the shape of the implied inflation rate is quite similar to the actual inflation rate, and they both peak in mid-2022. Furthermore, even though I only consider oil supply shocks and abstract from all other potential exogenous drivers of inflation, the implied inflation rate increases from the initial value of 4.1% up to 7%. Thus, the oil supply shocks account for about 60% of the additional increase from early 2021 until mid-2022. This is consistent with the findings in Shapiro (2023) who estimates the supply-driven part of (PCE) inflation in 2022 to be of similar magnitude.

Figure 5: Supply shocks, attention and the post-Covid inflation surge



Notes: This figure shows the actual US CPI inflation from 2020 until 2023, as well as the counterfactual inflation dynamics arising only from oil supply news shocks in the economy accounting for the increase in attention in April 2021 (blue-dashed line) and in the economy ignoring the increase in attention (orange-dotted line).

How much of this is due to the increase in attention? The red-dotted line shows that if there

[16] A similar approach is, e.g., used in Coibion (2012) who analyzes the historical contribution of monetary policy shocks to industrial production, unemployment and inflation, or by Mitra (2023) who examines the role of noise shocks for the dynamics of labor markets.

[17] I ignore all previous shocks, i.e., I compute the counterfactuals as if the economy was in steady state in April 2021 and that there were no shocks before that period. Thus, the reported results in this section are likely to be at the lower end of the actual importance of attention for the inflation dynamics, given that inflation was already increasing before and that the supply shocks in the first three months of 2021 were all negative (i.e., inflationary). Consistent with early 2021 being the time attention increased, Hilscher et al. (2022) document a rise in the probability of persistent high inflation for the US in mid-2021.

would not have been an increase in people's attention, that is when relying on the estimated impulse response functions from the low-attention regime (panel (b) in Figure 2), the implied inflation increase from April 2021 until mid-2022 would have only been about 1.4 percentage points (compared to the 2.9pp. when accounting for people's change in attention). These results suggest that the rise in the public's attention to inflation was quantitatively important in driving inflation further up in response to supply shocks during the recent inflation surge, and that without the increase in people's attention supply-driven inflation would have only been about half as high as it was. Using the estimates from Figure 3, i.e., when using Google searches as the threshold-defining variable, I find that supply shocks in the recent inflation surge likely contributed even more than implied by Figure 5, and the role of attention is even bigger (see Figure 26 in Appendix C.3).

To better understand the underlying mechanisms of this section's results, I now introduce the attention threshold and changing degrees of people's attention to inflation into an otherwise standard New Keynesian model.

4 A Monetary Model with Varying Attention

This section presents the monetary model which builds on the textbook New Keynesian model (Woodford, 2003; Galí, 2015) but households and firms pay only limited attention to inflation and the output gap.

4.1 Households

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

$$\tilde{E}_0 \sum_{t=0}^{\infty} \beta^t Z_t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Xi \frac{H_t^{1+\nu}}{1+\nu} \right], \tag{7}$$

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. Ξ is the utility weight on hours worked.

Households maximize their lifetime utility subject to the flow budget constraints

$$C_t + B_t = w_t H_t + \frac{1+i_{t-1}}{1+\pi_t} B_{t-1} + \frac{T_t}{P_t}, \qquad \text{for all } t, \tag{8}$$

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 denotes lump-sum taxes and transfers from the government.

Maximizing (7) subject to (8) yields the Euler equation

$$Z_t C_t^{-\sigma} = \beta(1+i_t)\tilde{E}_t \left[Z_{t+1} C_{t+1}^{-\sigma} \frac{1}{1+\pi_{t+1}} \right], \tag{9}$$

and the labor-leisure condition

$$w_t C_t^{-\sigma} = \Xi H_t^{\nu}. \tag{10}$$

4.2 Firms

The firm sector is held by risk-neutral managers that discount future profits by β and they have a mass of zero, such that their consumption is 0 and all their profits go to the households, as in Bayer et al. (2022).

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

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{11}$$

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:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t. \tag{12}$$

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

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}. \tag{13}$$

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

$$Y_t(j) = H_t(j). \tag{14}$$

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.[18] These taxes are given back to firms in a

[18] Therefore, we have $1 - \tau = \frac{\epsilon}{\epsilon-1}$ in the steady state.

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 one sets the price of firm j given these forecasts, similar to the setup in, e.g., Adam and Padula (2011). I discuss here the problem of the price setter and discuss the forecaster's problem later.

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

$$(1-\tau_t)P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}\frac{Y_t}{P_t} - w_t H_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 Y_t + t_t^F(j), \tag{15}$$

where $\psi \geq 0$ captures the price-adjustment cost parameter. The price setters sets the price to maximize profits

$$\Omega_0(j) = \tilde{E}_0^j \sum_{t=0}^{\infty} \beta^t \left[(1-\tau_t)P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}\frac{Y_t}{P_t} - mc_t H_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 Y_t + t_t^F(j)\right],$$

where $mc_t = w_t$ 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

$$\begin{aligned}(1-\tau_t)(\epsilon-1)P_t(j)^{-\epsilon}\frac{Y_t}{P_t^{1-\epsilon}} =& \epsilon mc_t\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon-1}\frac{Y_t}{P_t} - \psi\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)\frac{Y_t}{P_{t-1}(j)} \\ &+ \beta\psi\tilde{E}_t^j\left[\left(\frac{P_{t+1}(j)}{P_t(j)}-1\right)\frac{P_{t+1}(j)}{P_t(j)}\frac{Y_{t+1}}{P_t(j)}\right].\end{aligned}$$

Defining $T_t \equiv 1-\tau_t$, it follows that after a linearization around the zero-inflation steady state, firm j sets its price according to

$$\widehat{p}_t(j) = \frac{1}{\psi+\epsilon}\left[\psi\widehat{p}_{t-1} + \epsilon\left(\widehat{mc}_t - \widehat{T}_t + \widehat{p}_t\right) + \beta\psi\tilde{E}_t^j\pi_{t+1}^j\right], \tag{16}$$

where hatted variables denote log deviations of the respective variables from their steady state values (see Appendix D for all derivations). Therefore, prices may only differ across firms j due to differences in forecasts $\tilde{E}_t^j\pi_{t+1}^j$.

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 to households and $t_t^F(j)$ to firms.

The real government budget constraint is given by

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

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.

The monetary authority sets the nominal interest rate, according to a (linearized) Taylor rule:

$$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1-\rho_i)\left(\phi_\pi \pi_t + \phi_x \widehat{x}_t\right), \tag{17}$$

where $\tilde{i}_t$ denotes the nominal interest rate in deviations from its steady state, ρ_i captures interest-rate smoothing, ϕ_π and ϕ_x pin down the response coefficients with respect to inflation and the output gap, respectively. I discuss other rules for monetary policy in Section 6.

4.3 Subjective expectations under limited attention

Households. Households believe that consumption and inflation both follow a random walk:

$$\pi_t = \pi_{t-1} + \nu_{\pi,t}$$
$$\widehat{c}_t = \widehat{c}_{t-1} + \nu_{c,t},$$

where $\nu_{\pi,t}$ and $\nu_{c,t}$ are normally distributed with mean zero, time-invariant standard deviations and independent from each other. The random walk assumption ensures that long-run expectations align with the actual long-run variables.[19] They observe the aggregate shocks τ_t and Z_t perfectly.

At the time the household forms her expectations about future consumption and inflation, she does not perfectly observe their current realizations. This assumption could capture that the household head is not at all times perfectly monitoring all consumption expenditures of the members of the household and that inflation is not perfectly observable in real time. Instead, the household only sees noisy signals of the form

$$s_{\pi,t} = \pi_t + \varepsilon_{\pi,r,t}$$
$$s_{c,t} = \widehat{c}_t + \varepsilon_{c,t},$$

with normally distributed noise terms $\varepsilon_{\pi,r,t}$ and $\varepsilon_{c,t}$. Following the attention choice problem outlined in Section 2, the noise variance regarding inflation $\varepsilon_{\pi,r,t}$ is smaller in the H regime, reflecting the higher attention in that regime.

[19] Imposing a perceived law of motion following an AR(1) process does not qualitatively change the results.

As detailed in Appendix D, potential output, i.e., output under flexible prices, is constant and consumption in log-deviations from steady state is equal to the output gap in equilibrium, $\widehat{c}_t = \widehat{x}_t$.[20] It follows that if we assume initial values $\tilde{E}_{-1}\widehat{c}_0 = \tilde{E}_{-1}\widehat{x}_0$, we have $\tilde{E}_t\widehat{c}_{t+1} = \tilde{E}_t\widehat{x}_{t+1}$ for all t. This holds true, even if the household does not know that the two are equal in equilibrium.

Expectations about the output gap, $\widehat{x}_t$, then evolve as follows:

$$\tilde{E}_t\widehat{x}_{t+1} = \tilde{E}_{t-1}\widehat{x}_t + \gamma_x\left(\widehat{x}_t - \tilde{E}_{t-1}\widehat{x}_t\right), \tag{18}$$

where $\tilde{E}_t\widehat{x}_{t+1}$ denotes the agent's expectations of the one-period-ahead output gap. The parameter γ_x denotes the optimal level of attention to the output gap, based on the agent's subjective model of how the output gap evolves. A higher γ_x denotes a higher attention level. If $\gamma_x = 0$, the agent is completely inattentive and just sticks to her prior belief $\tilde{E}_{t-1}\widehat{x}_t$, whereas $\gamma_x = 1$ captures the case of full attention in which case the agent believes $\tilde{E}_t\widehat{x}_{t+1} = \widehat{x}_t$, which is the full-information belief of someone who believes that the output gap follows a random walk. As I discuss in more detail in the calibration section later on, I do not find any differences in attention to unemployment changes (which I use as a proxy for the output gap) in the data. I therefore impose that γ_x does not change across regimes.

Inflation expectations follow the law of motion derived in equation (2) for $\rho_\pi = 1$, so that they are given by

$$\tilde{E}_t\pi_{t+1} = \begin{cases} \tilde{E}_{t-1}\pi_t + \gamma_{\pi,L}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right), & \text{when } \pi_{t-1} < \bar{\pi} \\ \tilde{E}_{t-1}\pi_t + \gamma_{\pi,H}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right), & \text{when } \pi_{t-1} > \bar{\pi}, \end{cases} \tag{19}$$

where $\gamma_{\pi,r}$ captures the optimal level of attention to inflation in regime $r \in \{L, H\}$, depending on the respective cost of information $\frac{1}{\lambda_r}$. In what follows, I abstract from noise shocks and instead assume that the signal the household receives is exactly equal to the variable's realization in that period, but the household does not know this and acts as if there was noise.

Firm managers. Since there are no idiosyncratic shocks, I assume that the forecasting manager of firm j uses expectations about aggregate inflation to form her expectations about firm j's future price change, i.e., $\tilde{E}_t^j\pi_{t+1}(j) = \tilde{E}_t^j\pi_{t+1}$. Forecasting managers hold the same subjective expectations as households, consistent with McClure et al. (2022) who show that managers and non-managers hold similar average inflation and unemployment expectations and respond similarly to information treatments.[21] As for households, I abstract from noise shocks, which implies that all forecasters receive the same signal about current inflation, from

[20] Note, that the price adjustment costs do neither affect the steady state nor the linearized resource constraint when linearized around the zero-inflation steady state, such that $\widehat{y}_t = \widehat{c}_t$.

[21] McClure et al. (2022) further show that managers' expectations indeed affect their economic decisions.

which it follows that $\tilde{E}_t^j \pi_{t+1}(j) = \tilde{E}_t^j \pi_{t+1} = \tilde{E}_t \pi_{t+1}$.

These assumptions imply that all firms set the same price (as they do under rational expectations). Thus, firm-specific inflation is indeed equal to aggregate inflation, which confirms the forecaster's belief that the two are equal, and hence, the forecaster has no reason to deviate from these beliefs.

Given these assumptions, equation (16) can be written as

$$\pi_t = \frac{\epsilon}{\psi}\left(\widehat{mc}_t - \widehat{T}_t\right) + \beta \tilde{E}_t \pi_{t+1}. \tag{20}$$

From the labor-leisure equation and the production function, we have $\widehat{mc}_t = (\sigma + \nu)\,\widehat{y}_t$, and since potential output is constant we have $\widehat{y}_t = \widehat{x}_t$. Defining cost-push shocks as $u_t \equiv -\frac{\epsilon}{\psi}\widehat{T}_t$ and $\kappa \equiv \frac{\epsilon}{\psi}(\sigma + \nu)$, we arrive at the linearized New Keynesian Phillips Curve:

$$\pi_t = \beta \tilde{E}_t \pi_{t+1} + \kappa \widehat{x}_t + u_t. \tag{21}$$

4.4 Equilibrium

The model can then be summarized by three equilibrium equations when expressed in log-deviations from the zero-inflation steady state (see Appendix D):

$$\pi_t = \beta \tilde{E}_t \pi_{t+1} + \kappa \widehat{x}_t + u_t, \tag{22}$$

$$\widehat{x}_t = \tilde{E}_t \widehat{x}_{t+1} - \varphi\left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - r_t^*\right), \tag{23}$$

$$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1 - \rho_i)\left(\phi_\pi \pi_t + \phi_x \widehat{x}_t\right), \tag{24}$$

as well as the two law of motions for output gap expectations and inflation expectations, equations (18) and (19). Equation (22) is the New Keynesian Phillips curve, representing the supply side of the economy, and equation (23) denotes the aggregate Euler (or IS) equation, which together with monetary policy (equation (24)) pins down aggregate demand. $\varphi > 0$ measures the real rate elasticity of output, and r_t^* is the natural interest rate. The natural interest rate is the real rate that prevails in the economy with fully flexible prices, and solely depends on the exogenous shocks Z_t. 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. I will refer to shocks to r_t^* as demand shocks. The nominal interest rate $\tilde{i}_t$ and the natural rate are both expressed in absolute deviations of their respective steady state values, $\underline{i}$ and $\underline{r}^*$, with $\underline{i} = \underline{r}^*$, as the model is linearized around the zero-inflation steady state.

5 Inflation Surges

In this section, I show how the model with two inflation-attention regimes can jointly generate persistent heightened inflation periods, forecast error dynamics mirroring the ones recently observed in the United States, and lead to an asymmetry in the dynamics of inflation. Before going into these numerical results, however, I use a stylized version of my model economy to illustrate how the attention threshold alters the inflation dynamics and how shocks become more inflationary in times of high attention.

5.1 An Illustrative Example

To provide intuition how the attention threshold can trigger self-reinforcing inflation surges, I start with a slightly stylized version of the model. In particular, I assume that agents are completely inattentive to the output gap, i.e., $\gamma_x = 0$, and that the Taylor rule is given by $\tilde{i}_t = \phi_\pi \pi_t$ with $\phi_\pi > 1$. For readability, I set the interest-rate elasticity φ to 1.

The economy starts in the steady state, with $u_0 = 0$, $r_0^* = 0$ and prior expectations are at their long-run averages of 0: $\tilde{E}_{-1}\pi_0 = 0$ and $\tilde{E}_{-1}\widehat{x}_0 = 0$. The aggregate supply (AS) equation—captured by the Phillips Curve—in this initial period is then given by

$$AS_0: \quad \pi_0 = \frac{\kappa}{1-\beta\gamma_{\pi,L}}\widehat{x}_0, \tag{25}$$

and aggregate demand (AD), which follows from combining the Taylor rule with the aggregate Euler (or IS) equation, is given by

$$AD_0: \quad \pi_0 = -\frac{1}{\phi_\pi - \gamma_{\pi,L}}\widehat{x}_0. \tag{26}$$

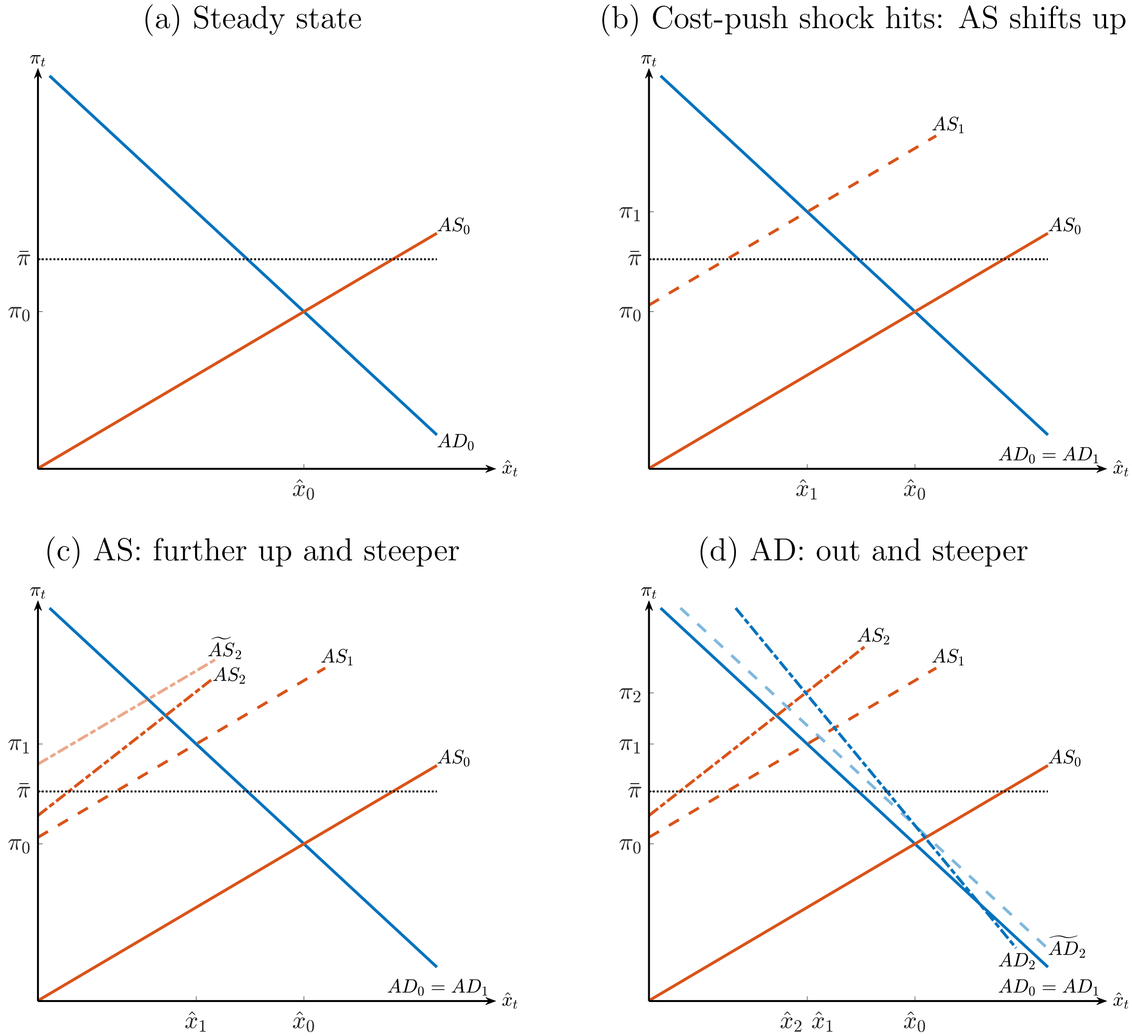
Panel (a) in Figure 6 depicts this initial situation graphically. Inflation and the output gap are both at their steady state values of 0 and therefore, below the inflation attention threshold $\bar{\pi}$.

In period 1, a positive cost-push shock hits and I assume that it persists for two periods: $u_1 = u_2 > 0$ and returns to zero afterwards, $u_t = 0$ for $t \geq 3$. The AS and AD equations are now given by

$$\begin{aligned} AS_1: \quad \pi_1 &= \frac{\kappa}{1-\beta\gamma_{\pi,L}}\widehat{x}_1 + \frac{1}{1-\beta\gamma_{\pi,L}}u_1 \\ AD_1: \quad \pi_1 &= -\frac{1}{\phi_\pi - \gamma_{\pi,L}}\widehat{x}_1. \end{aligned}$$

This situation is shown in panel (b) of Figure 6. The cost-push shock shifts the AS curve up along the AD curve. The resulting equilibrium is characterized by output below potential, i.e., a negative output gap, and positive inflation. The shock is assumed to be large enough, such

Figure 6: An illustrative example



Notes: Panel (a) shows the initial situation in period 0 when the economy is in the steady state. In period 1, a positive cost-push shock hits the economy, leading to an upward shift of the AS curve (panel (b)). In period 2, the AS curve shifts further up (shift from AS_1 to $\widetilde{AS}_2$) and the curve becomes steeper (rotation from $\widetilde{AS}_2$ to AS_2), shown in panel (c). Simultaneously, the AD curve shifts out (shift from AD_1 to $\widetilde{AD}_2$) and becomes stepper (rotation from $\widetilde{AD}_2$ to AD_2), as shown in panel (d). The black-dotted line at $\bar{\pi}$ depicts the inflation attention threshold.

that inflation exceeds the threshold.

Due to the increase in inflation in period 1, firms enter the second period with positive prior inflation expectations: $\tilde{E}_1\pi_2 = \gamma_{\pi,L}\pi_1 > 0$. These higher prior expectations together with the still ongoing cost-push shock shift the AS curve further up. This shift is illustrated in panel (c) of Figure 6 by the $\widetilde{AS}_2$ curve, which is given by

$$\widetilde{AS}_2: \quad \pi_2 = \frac{\kappa}{1-\beta\gamma_{\pi,L}}\widehat{x}_2 + \underbrace{\frac{1}{1-\beta\gamma_{\pi,H}}u_2 + \frac{\beta(1-\gamma_{\pi,H})\gamma_{\pi,L}}{1-\beta\gamma_{\pi,H}}\pi_1}_{\text{Intercept}>0}.$$

The terms denoted "Intercept" capture this shift in the AS curve. Since inflation in the previous period exceeded the attention threshold, attention is now higher. This increase in attention leads to an unambiguously stronger effect of the cost-push shock compared to the case in which attention would have remained constant:

$$\frac{1}{1-\beta\gamma_{\pi,H}}u_2 > \frac{1}{1-\beta\gamma_{\pi,L}}u_2.$$

The cost-push shock leads to an increase in inflation, and this inflation increase now leads to a larger increase in firms' inflation expectations due to their higher attention. These higher expectations then feed back into higher prices and thus, higher inflation.

The effect of the increase in firm managers' prior expectations on inflation in the second period, however, is smaller at the higher attention level. There are two counteracting forces. First, the increase in attention means that firm managers now update their expectations more strongly and put less weight on their prior expectations. This per se leads to a smaller effect. Second, the higher prior increases overall inflation expectations which, ceteris paribus, increases current inflation. But because inflation expectations are discounted by $\beta \leq 1$, the first effect dominates. Therefore, the shift in the AS curve due to the higher prior expectations is smaller at higher attention levels. Quantitatively, however, these differences are very small as $\beta \approx 1$. In fact, if $\beta = 1$, the change in attention has no effect on the shift due to higher prior expectations, and therefore, the total shift of the AS curve is unambiguously higher when attention is higher, due to stronger effect of the cost-push shock.

The shift of the AS curve to $\widetilde{AS}_2$, however, is only part of the story. What I ignored so far is that the increase in attention also increases the slope of the AS curve. That is, the Phillips Curve becomes steeper in periods of high attention—a *dynamic* non-linearity.[22] Taking this into account, the AS curve in the second period is given by

$$AS_2: \quad \pi_2 = \underbrace{\frac{\kappa}{1-\beta\gamma_{\pi,H}}}_{\text{Slope}}\widehat{x}_2 + \frac{1}{1-\beta\gamma_{\pi,H}}u_2 + \frac{\beta(1-\gamma_{\pi,H})\gamma_{\pi,L}}{1-\beta\gamma_{\pi,H}}\pi_1.$$

This steepening of the AS curve is illustrated in panel (c) of Figure 6 by the rotation of the AS curve from $\widetilde{AS}_2$ to AS_2. This steepening of the AS curve eases the inflationary pressures due to the negative output gap. Nevertheless, the steeper AS curve implies that if the AD curve would now shift out, the inflationary effects of this increase in demand would become larger.

[22]This non-linearity is equally present in the model without the simplifying assumptions I made at the beginning of this subsection. Note, that this is a different form of non-linearity as for example in Benigno and Eggertsson (2023) where the Phillips curve becomes steeper at a higher level of labor market tightness, whereas in my case, the slope does not depend on the level of real activity but depends on people's attention. When attention is high, the Phillips curve becomes steeper at every level of real activity.

It turns out that the higher prior expectations lead to such a demand increase. This is illustrated by $\widetilde{AD}_2$ in panel (d) of Figure 6, which is given by

$$\widetilde{AD}_2: \quad \pi_2 = -\frac{1}{\phi_\pi - \gamma_{\pi,L}}\widehat{x}_2 + \frac{(1-\gamma_{\pi,H})\gamma_{\pi,L}}{\phi_\pi - \gamma_{\pi,H}}\pi_1.$$

The higher prior expectations, ceteris paribus, decrease the real rate which leads to the outward shift of the AD curve. These effects, however, are smaller at higher levels of attention as long as the Taylor principle, $\phi_\pi > 1$, is satisfied, because in that case the higher inflation rates due to the higher prior expectations are counteracted by a more than one-for-one increase in the nominal rate. If monetary policy is relatively dovish, i.e., ϕ_π is close to 1, these differences are small. Furthermore, since the AD curve is now shifted along a steeper AS curve due to the heightened attention, the inflationary effects of a given shift are larger at the higher level of attention. This also implies that additional demand stimulus—for example, due to loose monetary policy or a fiscal stimulus—would have relatively large inflationary effects.

Additionally, the AD curve also becomes steeper, as illustrated by the rotation from $\widetilde{AD}_2$ to AD_2 in panel (d), where AD_2 is given by

$$AD_2: \quad \pi_2 = -\frac{1}{\phi_\pi - \gamma_{\pi,H}}\widehat{x}_2 + \frac{(1-\gamma_{\pi,H})\gamma_{\pi,L}}{\phi_\pi - \gamma_{\pi,H}}\pi_1.$$

This leads to a further increase in inflation, especially now because the AS curve is steeper. Consistent with the empirical findings in Section 3, an adverse supply shock would now be more inflationary due to the steeper AD curve. In this stylized example, inflation increases substantially from period 1 to period 2 through the change in attention, whereas the output gap remains practically constant.[23] Thus, the attention threshold offers a theory of how inflation surges may occur and exhibit self-reinforcing dynamics without changes in output.

In the third period, when the shock has died out, the AS curve shifts back down. The AS curve is given by

$$AS_3: \quad \pi_3 = \frac{\kappa}{1-\beta\gamma_{\pi,H}}\widehat{x}_3 + \frac{\beta(1-\gamma_{\pi,H})}{1-\beta\gamma_{\pi,H}}\tilde{E}_2\pi_3.$$

Due to the positive prior expectations, $\tilde{E}_2\pi_3 > 0$, the AS curve does not fully shift back to its initial position but remains above it. Due to the higher steepness of the AD curve, however, the shift in the AS curve leads to a stronger reduction in inflation compared to a flatter AD curve.

[23] The additional inflation surge due to the attention shift is also noticeable when comparing the model with the one that does not feature the increase in attention. The inflation increase from period 1 to period 2 in the model with the attention threshold is about 50%, but only about 20% in the model absent the attention change. The output gap is almost identical in both cases. The values I used in this stylized example are: $\bar{\pi} = 4\%$, $\gamma_{\pi,L} = 0.2$, $\gamma_{\pi,H} = 0.4$, $\beta = 0.99$, $\kappa = 0.6$, $\phi_\pi = 1.05$ and $u_1 = u_2 = 10$.

While the AS curve comes back down, the AD curve shifts further out due to the positive prior expectations that agents have when going into period 3. The AD curve is given by

$$AD_3: \quad \pi_3 = -\frac{1}{\phi_\pi - \gamma_{\pi,H}}\widehat{x}_3 + \frac{1-\gamma_{\pi,H}}{\phi_\pi - \gamma_{\pi,H}}\tilde{E}_2\pi_3.$$

Thus, output recovers more strongly during this disinflationary period than what would be the case absent this shift in the AD curve, and disinflationn occurs more gradually. These results are shown graphically in Figure 27 in Appendix E.

5.2 Large vs. small shocks

Equipped with these theoretical insights, I now move to a numerical analysis of the attention threshold. To do so, I set $\gamma_{\pi,L} = 0.18$ and $\gamma_{\pi,H} = 0.36$, and the attention threshold to $\bar{\pi} = 4\%$ (annualized), as estimated in Section 2. To calibrate the attention parameter with respect to the output gap γ_x, I follow the same procedure as for inflation but focus on expectations about unemployment changes (see Appendix B.1). This results in $\gamma_x = 0.24$ when inflation is below 4%, and $\gamma_x = 0.25$ when inflation is above it. The two estimates are not statistically significantly different from each other. I therefore impose that γ_x does not change across regimes and set it to $\gamma_x = 0.25$.

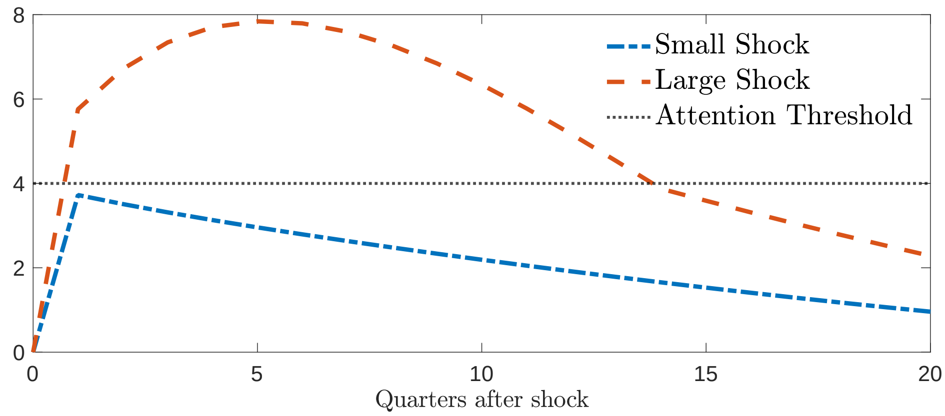
The rest of the calibration is standard. I set the discount factor β to target a steady state natural rate of 1% (annualized), the interest-rate sensitivity φ to 1, and the slope of the Phillips curve κ to 0.057. The Taylor rule coefficients are set to $\rho_i = 0.7$, $\phi_\pi = 2$ and $\phi_x = 0.125$.

Figure 7 shows the inflation dynamics following a cost-push shock that pushes inflation in the first period above the attention threshold $\bar{\pi}$ (the red-dashed line shows the inflation response, and the black-dotted line the attention threshold), as well as following a shock that does not push inflation above the threshold (blue-dashed-dotted line).[24]

The two cases exhibit fundamentally different inflation dynamics. When attention remains low, which happens in response to the relatively small shock, inflation increases on impact and then gradually returns back to its initial value of zero. After the larger shock, however, inflation keeps on increasing for about five periods before it peaks and then starts to decrease thereafter. While this decrease happens relatively fast at first, it slows down once inflation falls back below the attention threshold, and thus, inflation remains elevated quite persistently. The model under FIRE exhibits similar dynamics than the blue line but inflation returns even more quickly to its initial value (see Figure 29 in Appendix E).

[24] I set the shock persistence to 0.8. The dynamics following a demand shock look qualitatively similar and are shown in Figure 28 in Appendix E. Figure 30 shows the inflation response when attention to the output gap decreases when attention to inflation increases. The results are qualitatively similar, but the high-inflation period becomes even more persistent.

Figure 7: Inflation dynamics



To understand the dynamics in the model with the attention threshold, both regime switches that take place are key. The first regime switch occurs because the shock impulse is large enough to push inflation above the attention threshold. Thus, in the second period, agents become more attentive to inflation and hence, increase their inflation expectations more strongly in response to the forecast errors they make. For a given nominal rate, households now perceive the real rate to be lower and thus, increase their consumption in response. The attention regime change also matters for the supply side. Firms increase their inflation expectations more strongly which leads them to increase their prices more strongly. On top of that, the equilibrium inflation response to the cost-push shock in a given state of the economy is higher in the high-attention regime:

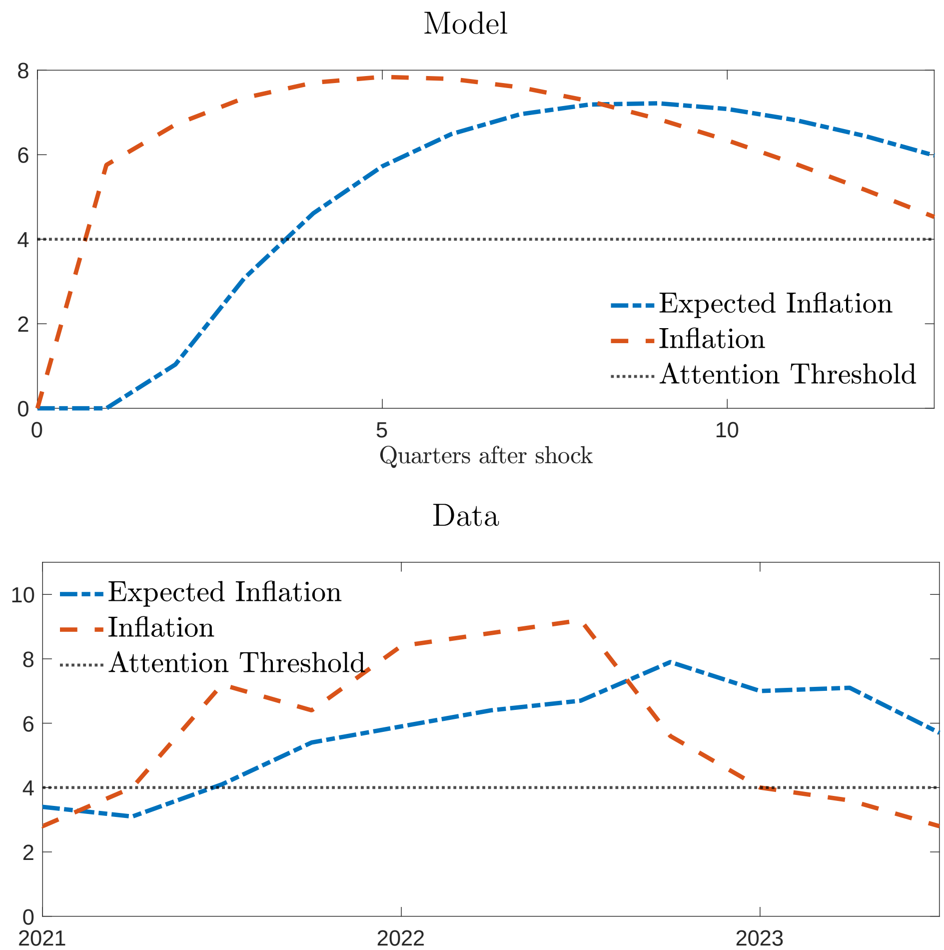
$$\frac{\partial \pi_{t+j}}{\partial u_{t+j}}\bigg|_{\gamma_{\pi,H}} > \frac{\partial \pi_{t+j}}{\partial u_{t+j}}\bigg|_{\gamma_{\pi,L}}.$$

Hence, inflation keeps on increasing, further fueling higher inflation expectations, leading to additional inflation increases. As the shock slowly dies out, inflation eventually starts to decline.

The second regime switch—once inflation falls back below the threshold value $\bar{\pi}$—leads to the slow inflation decrease after the initial inflation surge. When the economy enters the low-attention regime, agents decrease their attention to inflation. Thus, they mainly stick to their prior beliefs and only slowly update their expectations. As their priors are now relatively high due to the high inflation period, inflation expectations remain persistently high which hinders actual inflation from decreasing quickly.

To clearly see these patterns of inflation and inflation expectations, the upper panel in figure 8 plots expected and actual inflation, jointly. Inflation expectations are shifted such that the vertical distance between the two lines captures the forecast errors. Initially, expected inflation does not quite catch up with actual inflation, leading to positive forecast errors. After some time—around 8 quarters after the shock—when inflation is already decreasing, expected

Figure 8: Inflation and inflation expectation dynamics



Notes: The upper panel shows the dynamics of inflation and inflation expectations after a cost-push shock that pushes inflation above the attention threshold. The lower panel plots quarter-on-quarter (annualized) CPI inflation and one-quarter-ahead mean inflation expectations for the period 2021-2023, as well as the estimated attention threshold at 4%. In both panels, the inflation expectations are shifted such that the vertical distance between the two gives the respective forecast errors.

inflation surpasses actual inflation. Hence, forecast errors become negative.[25]

As the lower panel in figure 8 shows, these patterns are consistent with the recent inflation surge in the US (and more generally, also with the estimated impulse response functions of forecast errors in Figure 4). The figure shows annualized quarter-on-quarter CPI inflation and average inflation expectations from the Michigan Survey for the period from 2021 until 2023Q2. Consistent with the model implications, inflation peaks about a year and a half after its first increase. During this inflation surge, inflation expectations lagged actual inflation. As inflation peaked and began to decline, however, inflation expectations started to surpass actual inflation after about a year and a half after inflation exceeded the attention threshold. Consistent with

[25]Figures 31 and 32 in Appendix E show that the results look very similar when current rather than lagged inflation is the threshold-defining variable.

the model, inflation peaks at a higher value than inflation expectations. The model's dynamics are also consistent with the empirical findings in Blanco et al. (2022) who document, for a number of countries and different periods of inflation surges, that (i) inflation stays persistently high after the initial surge and (ii) that short-run inflation expectations initially fall short of actual inflation.

That inflation expectations stay persistently high, even when inflation has already fallen back below the threshold, gives rise to a heightened risk of another subsequent inflation surge. The higher inflation expectations keep actual inflation higher for longer, and therefore, closer to the attention threshold. Thus, a subsequent inflationary shock is more likely to push inflation back above the threshold and therefore, leading to another episode of persistently high inflation.

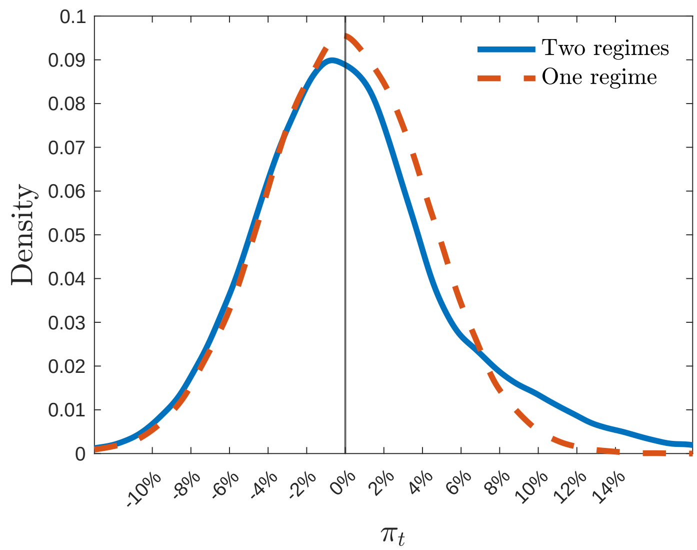
5.3 Asymmetry

The attention threshold induces an asymmetry in the dynamics of inflation. When inflation exceeds the threshold, attention increases, leading to a period of persistently high inflation. If, however, inflation is particularly low, attention remains unchanged. For example, when a shock pushes inflation from 0% to -4%, attention does not change. Therefore, we would not expect to see long-lived periods of pronounced deflation after such a shock. This matters for the overall properties of inflation. In particular, ignoring the attention threshold, would likely result in underpredicting the risk of persistently high inflation rates.

Figure 9 plots the distribution of inflation for the model with the attention threshold (blue-solid line) and for the model without the threshold (red-dashed line). These figures are obtained by simulating the model for 100,000 periods by hitting it with randomly-distributed cost-push shocks with zero mean and the same standard deviation as the oil supply news shocks from the empirical results in Section 3. I abstract from demand shocks for this exercise, but introducing demand shocks does not qualitatively change the insights in this section. The model with the attention threshold has a substantially thicker right tail than the one without the attention threshold. The probability that inflation exceeds 8% (annualized) is about 9% in the model with the threshold, and thus close to its empirical counterpart of 11%. In the model without the threshold, in contrast, inflation exceeds 8% only 3% of the time. When it comes to periods of low or moderate inflation, the two models exhibit very similar properties. For example, both models predict inflation to fall below -8% in only 3% of the periods, and the median in both models is very close to 0.[26] Thus, inflation in the model without the attention threshold is symmetric around its steady state value of 0, whereas it features a substantially higher risk of periods of high inflation in the model with the threshold.

[26]The mean is substantially above 0 in the model with the threshold, whereas it is practically 0 in the model without the threshold. I discuss this in more detail in Section 6 where I derive the normative implications of these results.

Figure 9: The attention threshold and inflation asymmetry



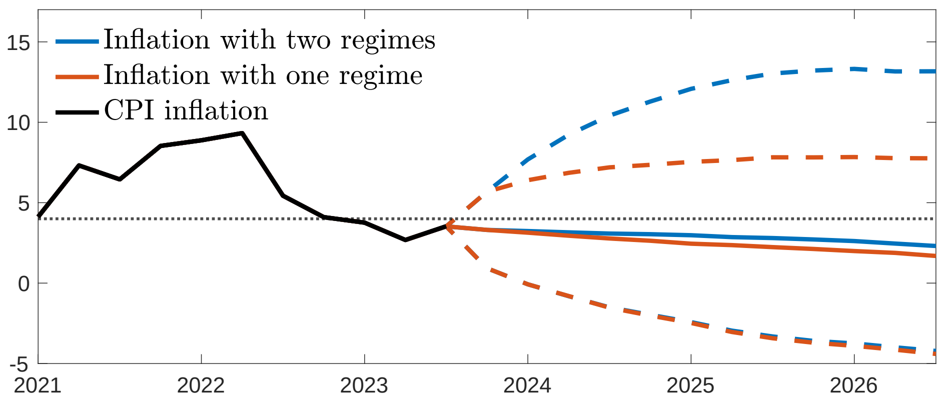
Notes: This figure shows the distribution of inflation (annualized) for the model with the attention threshold (solid-blue line) and the one without the attention threshold (dashed-red line).

This asymmetry matters for the current context and inflation outlook. To illustrate this, consider the following experiment: initialize the two model economies—one with and one without the attention threshold—at the most recent inflation number (the inflation rate in 2023Q3 was 3.53% annualized). I do this by assuming that there was one single supply shock in that month that pushed inflation up to 3.53%. Since this is below the threshold, both economies therefore start at exactly the same point. I then simulate the next three years 10,000 times for both models feeding in randomly-drawn supply shocks with zero mean and the same standard deviation as the one estimated for the oil supply news shocks used in the empirical section 3.

Figure 10 shows the results. The black line is the actually observed quarter-on-quarter annualized inflation from 2021Q1 until 2023Q3. The solid blue and red lines show the median dynamics of the 10,000 simulations for the economy with and without the attention threshold, respectively. The dotted lines depict the 5^{th} percentile and the 95^{th} percentile of the two model economies.

The model-implied median dynamics are such that in both economies, inflation slowly returns to its steady state value of 0. The two models also imply very similar dynamics at the lower end of the distribution of inflation dynamics. The 5^{th} percentile is in both cases slightly above -5% in 2026Q3. The risk of high inflation, however, differs significantly across the two models. While the 95^{th} percentile in the model without the attention threshold peaks slightly above an inflation rate of 7%, it increases to 13% in the model with the attention threshold. If negative supply shocks push the economy into the high-attention regime, the self-reinforcing

Figure 10: The attention threshold and the risk of high inflation



Notes: This figure shows the median inflation dynamics (solid lines) for the economy with the attention threshold (blue) and the one without the threshold (red) when simulating the model 10,000 times starting at the 2023Q3 inflation rate and randomly drawing normally distributed zero-mean cost-push shocks with a standard deviation as computed empirically for the oil supply news shocks. The dotted lines show the 5^{th} and 95^{th} percentiles of the simulations, respectively.

inflation dynamics discussed above would put upward pressure on inflation and we would experience a second period of persistently high inflation rates.

6 Welfare Implications

I now characterize the normative implications of the inflation attention threshold and the corresponding changes in attention. To compare different policy rules, I define welfare as

$$-\frac{1}{2}E_0\sum_{t=0}^{\infty}\beta^t\left[\pi_t^2+\Lambda\widehat{x}_t^2\right], \tag{27}$$

where Λ is the relative weight of the output gap, which I set to $\Lambda = 0.007$ as in Adam and Billi (2006).[27] When implementing the optimal policy, the policymaker maximizes welfare (27) subject to the private sector's optimality conditions, characterized by equations (22) and (23). I assume that the planner thinks that the private sector holds rational expectations. As Clarida et al. (1999) show, the optimal policy under commitment is given by

$$\pi_t + \tfrac{\Lambda}{\kappa}\left(\widehat{x}_t - \widehat{x}_{t-1}\right) = 0$$

[27] Welfare (27) can be obtained by deriving a second-order approximation to the household's utility function (see, e.g., Woodford (2003)). An implicit assumption here is that the policymaker is paternalistic (Benigno and Paciello, 2014) in the sense that the policymaker evaluates the household's utility under rational expectations.

when agents hold rational expectations. Under discretion, the optimal policy does not take its effects on inflation expectations into account and is then given by

$$\pi_t + \frac{\Lambda}{\kappa}\widehat{x}_t = 0.$$

In the following, I compare the welfare implications of these different policy rules and compare them to the Taylor rule (24) as well as to a Taylor rule without interest rate smoothing and no response to the output gap, i.e., to (24) where $\rho_i = \phi_x = 0$, as well as to a strict-inflation targeting regime in which inflation is kept at zero at all times. Table 4 summarizes these different policy rules.

Table 4: Monetary policy rules

Nr.	Name	Equation
(1)	Taylor rule with smoothing	$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1-\rho_i)(\phi_\pi \pi_t + \phi_x \widehat{x}_t)$
(2)	Taylor rule without smoothing	$\tilde{i}_t = \phi_\pi \pi_t$
(3)	Optimal RE commitment policy	$\pi_t + \frac{\Lambda}{\kappa}(\widehat{x}_t - \widehat{x}_{t-1}) = 0$
(4)	Optimal RE discretionary policy	$\pi_t + \frac{\Lambda}{\kappa}\widehat{x}_t = 0$
(5)	Strict inflation targeting	$\pi_t = 0$

I then simulate the economy for 10,000 periods for each of the different policy rules. As in section 5, I focus solely on cost-push shocks, with a persistence of 0.8 and $\sigma_u = 0.3\%$.[28] Allowing for demand shocks does not qualitatively affect the following results. Figure 11 plots welfare (27) for the model with the attention threshold (blue-solid line), with limited attention but absent the attention threshold (red-dashed line) and the model under FIRE (black-dashed-dotted line) for the 5 different policy rules. Panel (a) in Figure 12 shows the respective inflation volatilities, panel (b) the average level of inflation, and panel (c) the frequency of how often inflation is above the threshold of 4%.

There are two main takeaways from figures 11 and 12. First, following simple Taylor rules (policy rules (1) and (2)) leads to substantially larger welfare losses compared to optimal policy rules or a strict inflation targeting rule in the model with the attention threshold, especially for the case with interest-rate smoothing.[29] The reason is that the economy spends a substantial amount of time in the high-attention regime in which inflation is high and volatile (see panel (c) in figure 12). Due to the asymmetry of the attention threshold the average level of inflation is higher when the economy is in the high-attention regime frequently (panel (b)). This is

[28]To calibrate the standard deviation of the cost-push shocks at quarterly frequency, I sum up all monthly oil supply news shocks within a quarter and use the resulting standard deviation as my σ_u. I use the same approach for figures 9 and 10.

[29]Gáti (2022) shows that in her model with varying degrees of long-run-expectations anchoring, the welfare losses of Taylor rules with a low inflation-response coefficient may go to infinity, as in this case the model becomes unstable and inflation (and the output gap and interest rates) becomes explosive.

Figure 11: Welfare

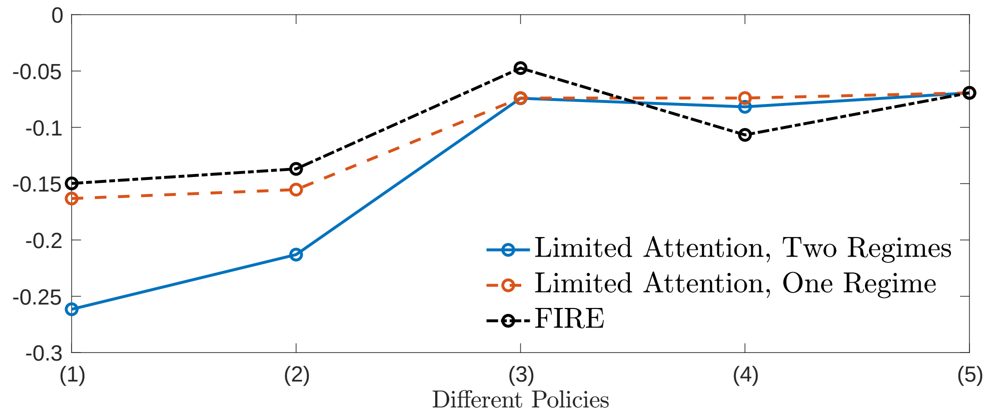
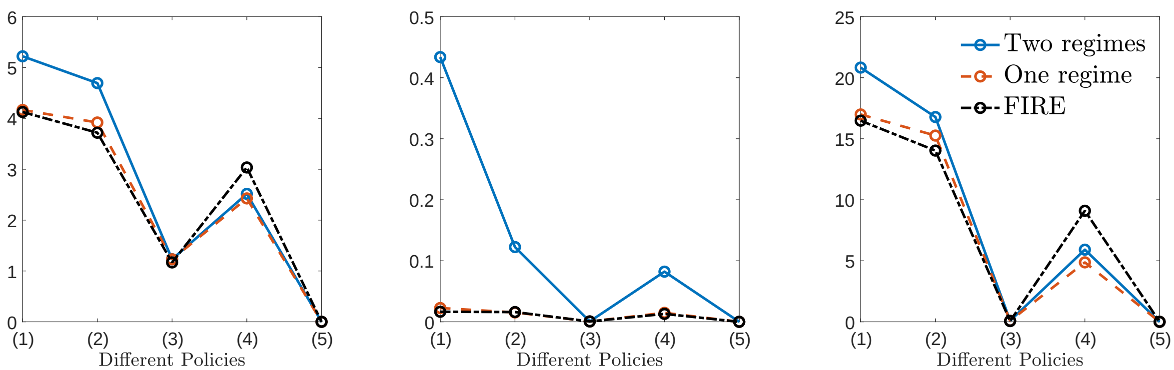


Figure 12: Inflation volatility, average inflation and time spent in the high-attention regime

(a) Inflation volatility (b) Average inflation (c) Frequency high-attention



in stark contrast to the model without the threshold or the one with fully-informed rational agents, where inflation fluctuates symmetrically around zero. Interest rate smoothing introduces additional persistence, such that the periods in the high-attention regime last longer and thus, the average level of inflation as well as its volatility increase.

The second main take away is that neither limited attention nor the attention threshold have large welfare implications, when monetary policy follows one of the rules (3)-(5). In all cases, inflation is relatively stable and fluctuates almost symmetrically around 0, as the economy very rarely stays in the high-attention regime (an exception is the optimal discretion rule (2)). In the case of the strict-inflation targeting regime (5) the inflation volatility is exactly 0. However, in that case, the output gap is more volatile (not shown) such that the overall welfare losses are very similar to the ones under the other policy rules from (3) and (4).

Overall, figures 11 and 12 illustrate that following simple Taylor rules, especially ones with interest-rate smoothing and relatively low inflation-response coefficients, can lead to large welfare losses when the public's attention to inflation may increase substantially during times of high inflation. Policies that are more hawkish and induce much smaller inflation fluctuations, in contrast, can mitigate the potentially detrimental effects of sudden increases in attention much more effectively, or even prevent these episodes from happening completely.

7 Conclusion

The recent inflation surge in many advanced economies brought inflation back on people's minds. In this paper, I quantify the inflation attention threshold after which people start to pay more attention to inflation. I estimate this attention threshold to be at an inflation rate of 4% in the United States and that attention doubles from the low-attention regime to the high-attention regime. Supply shocks become twice as inflationary in the high-attention regime and I find that the change in people's attention in early 2021 likely doubled the inflationary effect of supply shocks.

A New Keynesian model that accounts for the inflation attention threshold can replicate the empirical findings and produces inflation and inflation expectation dynamics that are consistent with the ones recently observed in the US.

Accounting for varying attention levels also matters for the models' normative predictions. Following simple Taylor rules leads to much larger welfare losses compared to optimal policy rules. Such simple policies induce frequent and long-lasting episodes in which attention is heightened which leads to large welfare losses, not only because these episodes increase the overall volatility of inflation, but—due to the asymmetry of the attention threshold—also the average inflation rates.

References

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.

Afrouzi, H. and C. Yang (2020): "Dynamic inattention, the phillips curve, and forward guidance," Tech. rep., Working Paper.

Amiti, M., S. Heise, F. Karahan, and A. Şahin (2023): "Inflation Strikes Back: The Role of Import Competition and the Labor Market," *Working paper.*

Andre, P., C. Pizzinelli, C. Roth, and J. Wohlfart (2022): "Subjective models of the macroeconomy: Evidence from experts and representative samples," *The Review of Economic Studies*, 89, 2958–2991.

Angeletos, G.-M., F. Collard, and H. Dellas (2020a): "Business-cycle anatomy," *American Economic Review*, 110, 3030–3070.

Angeletos, G.-M., Z. Huo, and K. Sastry (2020b): "Imperfect expectations: Theory and evidence," in *NBER Macroeconomics Annual 2020, volume 35*, University of Chicago Press.

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

Auerbach, A. J. and Y. Gorodnichenko (2012a): "Fiscal multipliers in recession and expansion," in *Fiscal policy after the financial crisis*, University of Chicago Press, 63–98.

——— (2012b): "Measuring the output responses to fiscal policy," *American Economic Journal: Economic Policy*, 4, 1–27.

Ball, L. M., D. Leigh, and P. Mishra (2022): "Understanding us inflation during the covid era," *Working paper.*

Bayer, C., B. Born, and R. Luetticke (2022): "Shocks, frictions, and inequality in US business cycles," .

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

Benigno, P. and G. B. Eggertsson (2023): "It's baaack: The surge in inflation in the 2020s and the return of the non-linear phillips curve," *Working paper.*

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

Bernanke, B. and O. Blanchard (2023): "What caused the US pandemic-era inflation?" *Hutchins Center Working Papers.*

Bhandari, A., J. Borovička, and P. Ho (2019): "Survey data and subjective beliefs in business cycle models," .

Bianchi, F., R. Faccini, and L. Melosi (2023): "A Fiscal Theory of Persistent Inflation," *The Quarterly Journal of Economics*, qjad027.

Bianchi, F. and L. Melosi (2022): "Inflation as a fiscal limit," .

Blanco, A., P. Ottonello, and T. Ranosova (2022): "The Dynamics of Large Inflation Surges," *Working paper.*

Bracha, A. and J. Tang (2023): "Inflation levels and (in) attention," *Working Paper*.

Broer, T., A. Kohlhas, K. Mitman, and K. Schlafmann (2021): "Information and wealth heterogeneity in the macroeconomy," *Working paper*.

Caggiano, G., E. Castelnuovo, and N. Groshenny (2014): "Uncertainty shocks and unemployment dynamics in US recessions," *Journal of Monetary Economics*, 67, 78–92.

Candia, B., O. Coibion, and Y. Gorodnichenko (2021): "The Inflation Expectations of US Firms: Evidence from a New Survey," *Working paper*.

Candia, B., M. Weber, Y. Gorodnichenko, and O. Coibion (2023): "Perceived and Expected Rates of Inflation of US Firms," 113, 52–55.

Carlson, J. A. and M. Parkin (1975): "Inflation expectations," *Economica*, 42, 123–138.

Carvalho, C., S. Eusepi, E. Moench, and B. Preston (2022): "Anchored inflation expectations," *American Economic Journal: Macroeconomics (forthcoming)*.

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. Gali, and M. Gertler (1999): "The science of monetary policy: a new Keynesian perspective," *Journal of economic literature*, 37, 1661–1707.

Coibion, O. (2012): "Are the effects of monetary policy shocks big or small?" *American Economic Journal: Macroeconomics*, 4, 1–32.

Coibion, O. and Y. Gorodnichenko (2015): "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. I. Knotek, and R. Schoenle (2020): "Average Inflation Targeting and Household Expectations," *Journal of Political Economy (forthcoming)*.

D'Acunto, F., U. Malmendier, and M. Weber (2023): "What do the data tell us about inflation expectations?" in *Handbook of economic expectations*, Elsevier, 133–161.

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2019): "Cognitive abilities and inflation expectations," in *AEA Papers and Proceedings*, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, vol. 109, 562–566.

D'Acunto, F., U. Malmendier, and M. Weber (2021): "Gender roles produce divergent economic expectations," *Proceedings of the National Academy of Sciences*, 118, e2008534118.

Evans, G. W. and G. Ramey (1995): "Expectation calculation, hyperinflation and currency collapse," *The new macroeconomics: Imperfect markets and policy effectiveness*, 307.

Flynn, J. P. and K. Sastry (2022): "Attention cycles," *Working Paper.*

Gagliardone, L. and M. Gertler (2023): "Oil Prices, Monetary Policy and Inflation Surges," *Working paper.*

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

Gallegos, J. E. (2023): "Inflation persistence, noisy information and the Phillips curve," *Working paper.*

Gáti, L. (2022): "Monetary Policy & Anchored Expectations An Endogenous Gain Learning Model," *Journal of Monetary Economics (forthcoming).*

Goldstein, N. (2023): "Tracking inattention," *Journal of the European Economic Association*, jvad022.

Gonzalo, J. and J.-Y. Pitarakis (2002): "Estimation and model selection based inference in single and multiple threshold models," *Journal of econometrics*, 110, 319–352.

Hansen, B. E. (2011): "Threshold autoregression in economics," *Statistics and its Interface*, 4, 123–127.

Hazell, J., J. Herreno, E. Nakamura, and J. Steinsson (2022): "The slope of the Phillips Curve: evidence from US states," *The Quarterly Journal of Economics*, 137, 1299–1344.

Hilscher, J., A. Raviv, and R. Reis (2022): "How likely is an inflation disaster?" .

Jarociński, M. and P. Karadi (2020): "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American economic review*, 95, 161–182.

Jørgensen, P. L. and K. J. Lansing (2023): "Anchored inflation expectations and the slope of the phillips curve," *Working Paper.*

Joussier, R. L., J. Martin, and I. Mejean (2023): "Energy cost pass-through and the rise of inflation: Evidence from French manufacturing firms," .

Känzig, D. R. (2021): "The macroeconomic effects of oil supply news: Evidence from OPEC announcements," *American Economic Review*, 111, 1092–1125.

Kim, G. and C. Binder (2023): "Learning-through-survey in inflation expectations," *American Economic Journal: Macroeconomics*, 15, 254–278.

Korenok, O., D. Munro, and J. Chen (2022): "Inflation and attention thresholds," *Working paper*.

Kroner, N. (2023): "Inflation and Attention: Evidence from the Market Reaction to Macro Announcements," *Available at SSRN 4527424*.

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.

Larsen, V. H., L. A. Thorsrud, and J. Zhulanova (2021): "News-driven inflation expectations and information rigidities," *Journal of Monetary Economics*, 117, 507–520.

Macaulay, A. (2022): "Shock transmission and the sources of heterogeneous expectations," *Working Paper*.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2018): "Dynamic rational inattention: Analytical results," *Journal of Economic Theory*, 176, 650–692.

——— (2020): "Rational Inattention: A Review," .

Mackowiak, B., F. Matejka, and M. Wiederholt (2023): "Rational Inattention: A Review," *Journal of Economic Literature*.

Mackowiak, B. and M. Wiederholt (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99, 769–803.

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.

McClure, E. M., O. Coibion, and Y. Gorodnichenko (2022): "The Macroeconomic Expectations of US Managers," *Working Paper*.

Meichtry, P. (2022): "Sticky Information, Heterogeneity, and Aggregate Demand," .

MITRA, A. (2023): "Imperfect Information and Slow Recoveries in the Labor Market," *Working paper.*

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

NIMARK, K. P. AND S. PITSCHNER (2019): "News media and delegated information choice," *Journal of Economic Theory*, 181, 160–196.

NORD, L. (2022): "Who Cares about Inflation? Endogenous Expectation Formation of Heterogeneous Households," *Working paper.*

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

PEDEMONTE, M., H. TOMA, AND E. VERDUGO (2023): "Aggregate Implications of Heterogeneous Inflation Expectations: The Role of Individual Experience," .

PFÄUTI, O. (2023): "Inflation–who cares? Monetary Policy in Times of Low Attention," *Working paper.*

PFÄUTI, O. AND F. SEYRICH (2023): "A behavioral heterogeneous agent New Keynesian model," *Working paper.*

RAMEY, V. A. AND S. ZUBAIRY (2018): "Government spending multipliers in good times and in bad: evidence from US historical data," *Journal of political economy*, 126, 850–901.

REIS, R. (2006a): "Inattentive consumers," *Journal of monetary Economics*, 53, 1761–1800.

——— (2006b): "Inattentive producers," *The Review of Economic Studies*, 73, 793–821.

——— (2022): "Losing the inflation anchor," *Brookings Papers on Economic Activity*, 2021, 307–379.

ROTEMBERG, J. J. (1982): "Sticky prices in the United States," *Journal of political economy*, 90, 1187–1211.

ROTH, C., M. WIEDERHOLT, AND J. WOHLFART (2023): "The Effects of Monetary Policy: Theory with Measured Expectations," *Working paper.*

SCHMIDT, T., H. MÜLLER, J. RIEGER, T. SCHMIDT, AND C. JENTSCH (2023): "Inflation Perception and the Formation of Inflation Expectations," *Working paper.*

SHAPIRO, A. H. (2023): "Decomposing supply and demand driven inflation," *Working paper.*

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

——— (2010): "Rational inattention and monetary economics," in *Handbook of monetary economics*, Elsevier, vol. 3, 155–181.

Tenreyro, S. and G. Thwaites (2016): "Pushing on a string: US monetary policy is less powerful in recessions," *American Economic Journal: Macroeconomics*, 8, 43–74.

Vellekoop, N. and M. Wiederholt (2019): "Inflation expectations and choices of households," .

Weber, M., F. D'Acunto, Y. Gorodnichenko, and O. Coibion (2022): "The subjective inflation expectations of households and firms: Measurement, determinants, and implications," *Journal of Economic Perspectives*, 36, 157–184.

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," .

Woodford, M. (2003): *Interest and Prices*, Princeton University Press.

A A Limited-Attention Model of Inflation Expectations

In this section, I derive the expectations-formation process under limited attention used 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

$$\max_f \int U(\pi^e, \pi) f(\pi^e, \pi) d\pi d\pi^e - C(f) \tag{28}$$

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

where $g(\pi)$ is the agent's prior, which is assumed to be Gaussian; $\pi \sim N(\widehat{\pi}, \sigma_\pi^2)$. $C(.)$ 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:

$$C(f) = \frac{1}{\lambda} I(\pi; \pi^e) = \frac{1}{\lambda} \left(H(\pi) - E\left[H(\pi | \pi^e) \right] \right),$$

where $H(x) = -\int f(x) log(f(x)) dx$ is the entropy of x and $\frac{1}{\lambda}$ measures the cost of information.

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

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

where χ measures the stakes of making a mistake.[30]

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 $\varepsilon \sim i.i.N.(0, \sigma_\varepsilon^2)$.[31] The problem (28) now reads

$$\max_{\sigma^2_{\pi|s} \leq \sigma^2_\pi} E_\pi \left[E_s \left[-\chi \rho_\pi^2 \left(\pi - E[\pi|s]\right)^2\right]\right] - \frac{1}{\lambda} I(\pi; \pi^e) = \max_{\sigma^2_{\pi|s} \leq \sigma^2_\pi} \left(-\chi \rho_\pi^2 \sigma^2_{\pi|s} - \frac{\frac{1}{\lambda}}{2} log \frac{\sigma^2_\pi}{\sigma^2_{\pi|s}}\right). \tag{29}$$

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

$$\pi^e = \rho_\pi \left(1 - \gamma\right) \widehat{\pi} + \rho_\pi \gamma s, \tag{30}$$

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

An equivalent way of writing γ is

$$\gamma = \frac{\sigma^2_\pi}{\sigma^2_\pi + \sigma^2_\varepsilon}. \tag{31}$$

Now, since the agent *chooses* the level of attention, we can re-formulate (29) as

$$\max_{\gamma \in [0,1]} \left(-\chi \rho_\pi^2 (1 - \gamma) \sigma_\pi^2 - \frac{\frac{1}{\lambda}}{2} log \frac{1}{1 - \gamma}\right). \tag{32}$$

Solving this optimization problem (32) yields the *optimal* level of attention

$$\gamma_\pi = max \left\{0, 1 - \frac{\frac{1}{\lambda}}{2\chi \rho_\pi^2 \sigma_\pi^2}\right\}. \tag{33}$$

The expression (33) for the optimal attention level shows that when the cost of information $\frac{1}{\lambda}$ is lower, attention is higher.

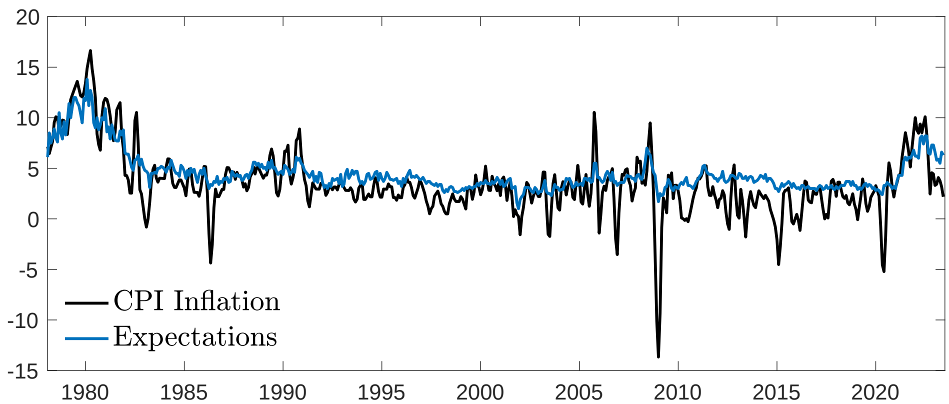
[30] A quadratic loss function is usually derived from a second-order approximation of the household's utility function or the firm's profit function (see, e.g., Mackowiak and Wiederholt (2009)).

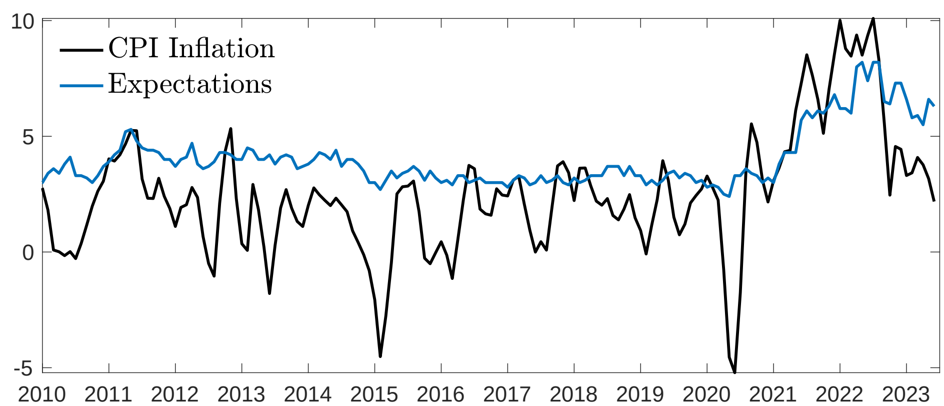
[31] In this case, the entropy becomes $H(x) = \frac{1}{2} log(2\pi e \sigma_x^2)$, where σ_x^2 is the variance of x. Note, that here π denotes the number "pi" and not inflation.

B Appendix to Empirical Section 2

Time series of inflation and inflation expectations.

Figure 13: Inflation and Inflation Expectations



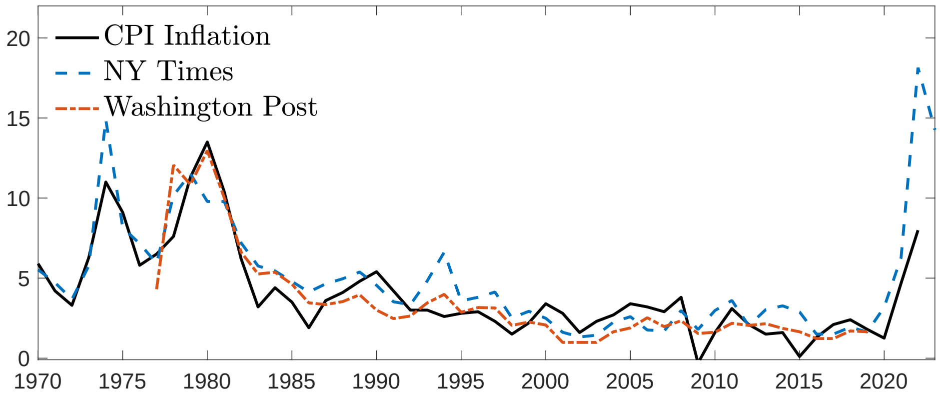


Notes: The black-solid lines show the annualized quarter-on-quarter CPI inflation rate at monthly frequency for the US, and the blue-dashed lines the average inflation expectations of households from the Survey of Consumers. The upper panel shows these time series for the period 1978 - 2023, and the lower panel for the period 2010 - 2023.

News Coverage of inflation. Figure 14 supports the assumption that media reporting about inflation is higher in times of higher inflation. The Figure shows the frequency of the word "inflation" in the New York Times (1970 to July 2023, blue-dashed line) and the Washington Post (1977-2019, red-dashed-dotted line), together with annual CPI inflation (black-solid line). There is a very strong positive correlation between inflation and news coverage of inflation (the correlation between CPI inflation and the two news-coverage series is 0.86 for the New York Times and 0.90 for the Washington Post).

News coverage of inflation is also substantially higher in times inflation is above the attention threshold. I find that the average frequency of the word "inflation" is 2.7 times as high for the New York Times and 2.9 times as high for the Washington Post when CPI inflation is above

Figure 14: News coverage of inflation is higher in times of high inflation



Notes: The black-solid line shows the annual CPI inflation rate for the US since 1970, the blue-dashed line and the red-dashed-dotted line show the frequency of the word inflation in the New York Times and the Washington post, respectively.

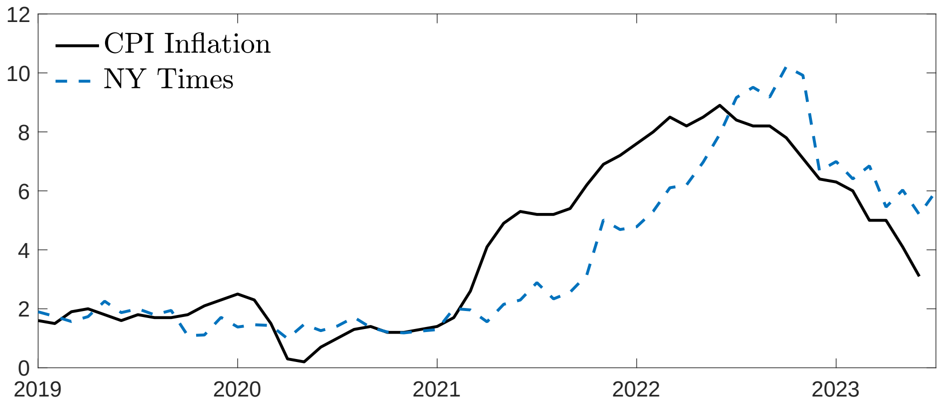
4% in that year compared to years in which CPI inflation is below 4%.

Consistent with this, Bracha and Tang (2023) show that higher inflation rates indeed lead to more media reporting about inflation, and Lamla and Lein (2014) find that more intensive news reporting about inflation improves the accuracy of consumers' inflation expectations, consistent with agents being more attentive. Schmidt et al. (2023) show that during episodes of intensive newspaper coverage of inflation, news reporting has strong effects on inflation expectations but not during other episodes. Larsen et al. (2021) find that news media coverage predicts households' inflation expectations, and Nimark and Pitschner (2019) show that major events (such as strong inflation increases) lead to a shift in the news focus towards these events.

Covid. When focusing on the sample from 2017 until 2023, I estimate an attention threshold of 3.63%, and attention levels of $\widehat{\gamma}_{\pi,L} = 0.03$ and $\widehat{\gamma}_{\pi,H} = 0.25$. The very low attention level before the outbreak of the Covid-19 pandemic confirms the findings in Pfäuti (2023) that attention to inflation was at a historical low at the time. When imposing the threshold at 4% (rather than estimating it), attention in the higher regime equals 0.28 and 0.04 in the low regime.

Figure 15 shows news coverage of inflation in the New York Times for the period 2019 until July 2023 at monthly frequency (blue-dashed line) together with monthly year-on-year CPI inflation. The figure shows the strong positive correlation of 0.85 between inflation and news coverage, and that at the peak, news coverage of inflation quintupled from its pre-pandemic level. Consistent with the theory, news coverage lags inflation slightly (the correlation between news coverage and *lagged* inflation is slightly higher than the one with current inflation, 0.9 instead of 0.85). The correlation of news coverage with Google searches is 0.94 for the period 2017-2023, further supporting the assumption that news coverage is highly correlated with the public's attention to inflation.

Figure 15: News coverage of inflation during Covid



Notes: The black-solid line shows the monthly year-on-year CPI inflation rate for the US and the blue-dashed line the frequency of the word inflation in the New York Times, normalized to have the same standard deviation as CPI inflation.

Heterogeneity. So far, I mainly focused on average (and median) expectations. I now test whether the attention threshold and the different attention levels depend on people's gender or their age.[32] As documented in D'Acunto et al. (2021), gender and gender roles (e.g., with respect to grocery shopping) play a big role in explaining differences in how men and women form their inflation expectations. I find that men have a higher attention threshold than women (4.4% vs. 3.9%), indicating that women increase their attention somewhat earlier than men. This might be explained by the fact that women are more likely to go grocery shopping than men (D'Acunto et al., 2021) and therefore, experience price changes more directly than men. I further find that the threshold for younger people (aged 18-34) is lower than for older people (4.44% vs. 6.8% for people aged between 35 and 54, and 5.7% for people aged older than 55), but that their attention levels tend to be lower overall (0.21 vs. 0.24 below the threshold and 0.41 vs. 0.74 above it).

B.1 Unemployment expectations

In order to calibrate the attention parameter with respect to the output gap, γ_x, I use expectations about unemployment changes from the Survey of Consumers. The survey asks households whether they expect unemployment to increase, decrease or to remain about the same over the next twelve months. I follow Carlson and Parkin (1975), Bhandari et al. (2019) and Pfäuti and Seyrich (2023) to translate these categorical unemployment expectations into numerical expectations.

Let q_t^D, q_t^S and q_t^U denote the shares reported at time t that think unemployment will go

[32] There is a vast literature documenting heterogeneity in inflation expectations, see, e.g., D'Acunto et al. (2019); Broer et al. (2021); Pfäuti and Seyrich (2023); D'Acunto et al. (2023); Pedemonte et al. (2023); Weber et al. (2022); Roth et al. (2023); Nord (2022); Meichtry (2022), for recent contributions.

down, stay roughly the same, or go up over the next year, respectively. I assume that these shares are drawn from a cross-sectional distribution of responses that are normally distributed according to $\mathcal{N}\left(\mu_t, (\sigma_t)^2\right)$ and a threshold a such that when a household expects unemployment to remain within the range $[-a, a]$ over the next year, she responds that unemployment will remain "about the same". We thus have

$$q_t^D = \Phi\left(\frac{-a - \mu_t}{\sigma_t}\right) \qquad q_t^U = 1 - \Phi\left(\frac{a - \mu_t}{\sigma_t}\right),$$

which after some rearranging yields

$$\begin{aligned}\sigma_t &= \frac{2a}{\Phi^{-1}\left(1 - q_t^U\right) - \Phi^{-1}\left(q_t^D\right)} \\ \mu_t &= a - \sigma_t \Phi^{-1}\left(1 - q_t^U\right).\end{aligned}$$

This leaves us with one degree of freedom, namely a. I follow Pfäuti and Seyrich (2023) and set $a = 0.5$ which means that if a household expects the change in unemployment to be less than half a percentage point (in absolute terms), she reports that she expects unemployment to be about the same as it is at the time of the survey. I use data from FRED for the actual unemployment changes and restrict the sample to end in 2019Q4, due to the extreme behavior of unemployment changes with the outbreak of the Covid-19 pandemic.

I then estimate γ_x separately for whether lagged inflation is above or below the estimated threshold of 4%. This results in estimates $\widehat{\gamma}_{x,L} = 0.25$ and $\widehat{\gamma}_{x,H} = 0.25$. Thus, there are no differences in attention to unemployment changes across regimes and hence, I impose γ_x to be the same across regimes in the model.

C Additional Results and Robustness Checks to Section 3

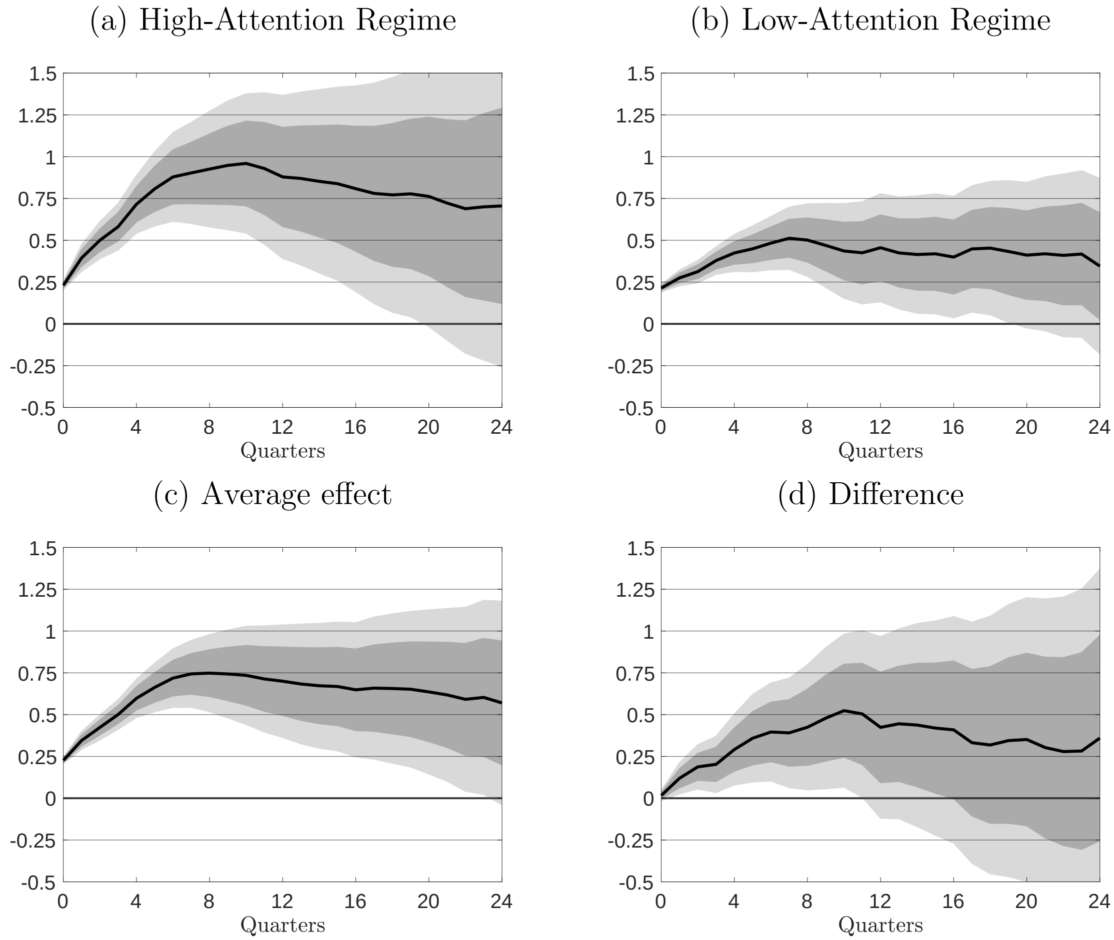
C.1 Other shocks

Inflation shock. Angeletos et al. (2020a) estimate a VAR including 10 key macroeconomic variables and then identify different shocks by maximizing its contribution to the volatility of a given variable over business-cycle frequency (6-32 quarters). I use their shock that contributes most to the volatility of inflation (using the GDP deflator, as in Angeletos et al. (2020a)). These shocks are available at quarterly frequency and span the period 1960 until the end of 2017. I use the previous quarter's CPI inflation as an indicator whether the economy is in the high-attention regime or the low-attention regime (i.e., was the previous annualized CPI inflation rate above or below 4%.)

The dependent variable is the change in the log of the GDP deflator (times 100) from quarter $t-1$ to quarter $t+h$. Figure 16 shows the results. As in the specification with oil news shocks, inflation responds about twice as much to the *inflation shock* when hit in the high-attention regime compared to the low-attention regime. As panel (d) shows, these differences are highly statistically significant and quite persistent: the difference is largest 10 quarters after the shock and still highly statistically significan.

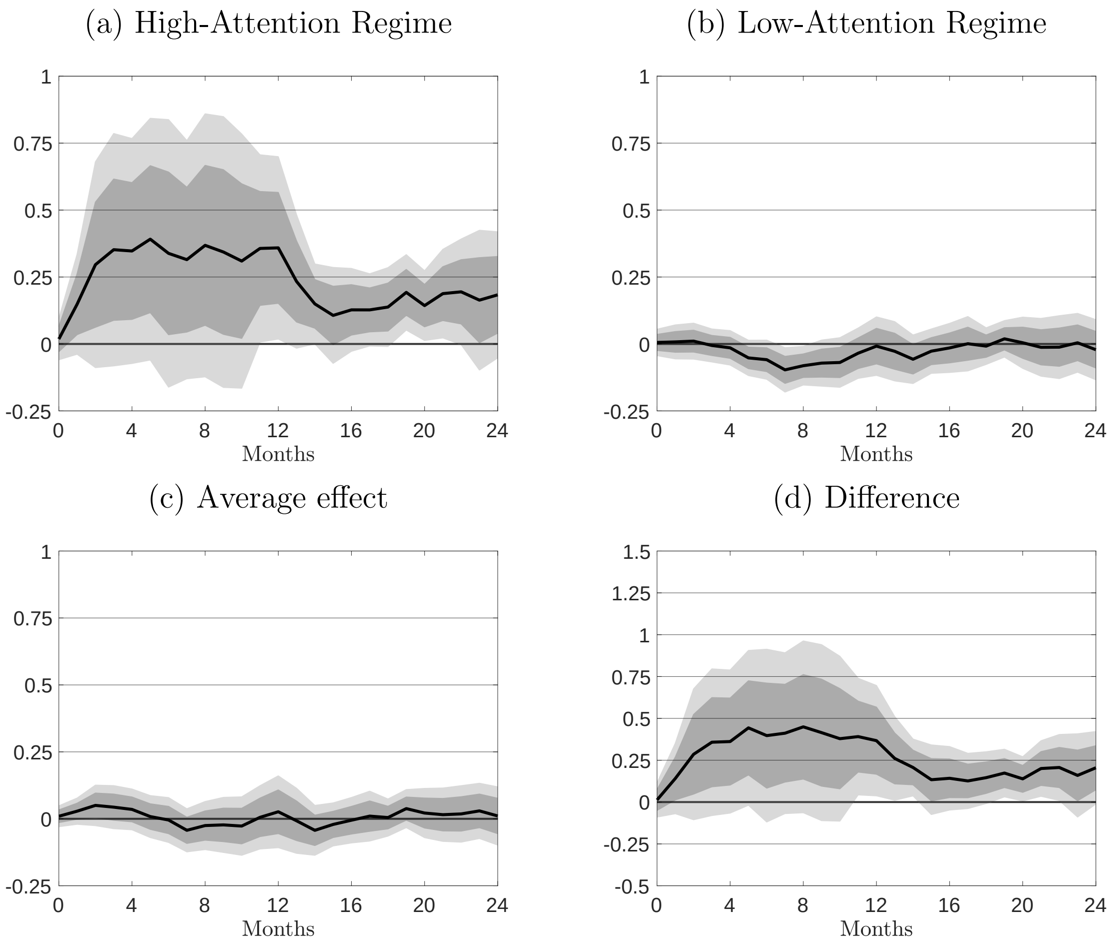
Monetary policy shocks. In figure 17, I show that inflation also responds more strongly in the high-attention regime to monetary policy shocks identified using a high-frequency identification (the shocks are taken from Jarociński and Karadi (2020) and are purged from the information effects of monetary policy statements; I show here the shocks identified using sign restrictions, but the results are practically identical when using the shock series based on the "poor-man approach"). A drawback of using these monetary policy shocks, however, is that they are only available for the period 1990-2019. Thus, the sample does not include most of the high-inflation periods. Therefore, while the differences across regimes are substantial in magnitude, the differences are less statistically significant.

Figure 16: Price level response to the shock targeting inflation



Notes: This figure shows the cumulative price level (using the GDP deflator) response to the shock of Angeletos et al. (2020a) that targets inflation in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 4 lags). The attention regimes ares defined based on the previous quarter's CPI inflation rate.

Figure 17: Inflation response to an monetary policy shock

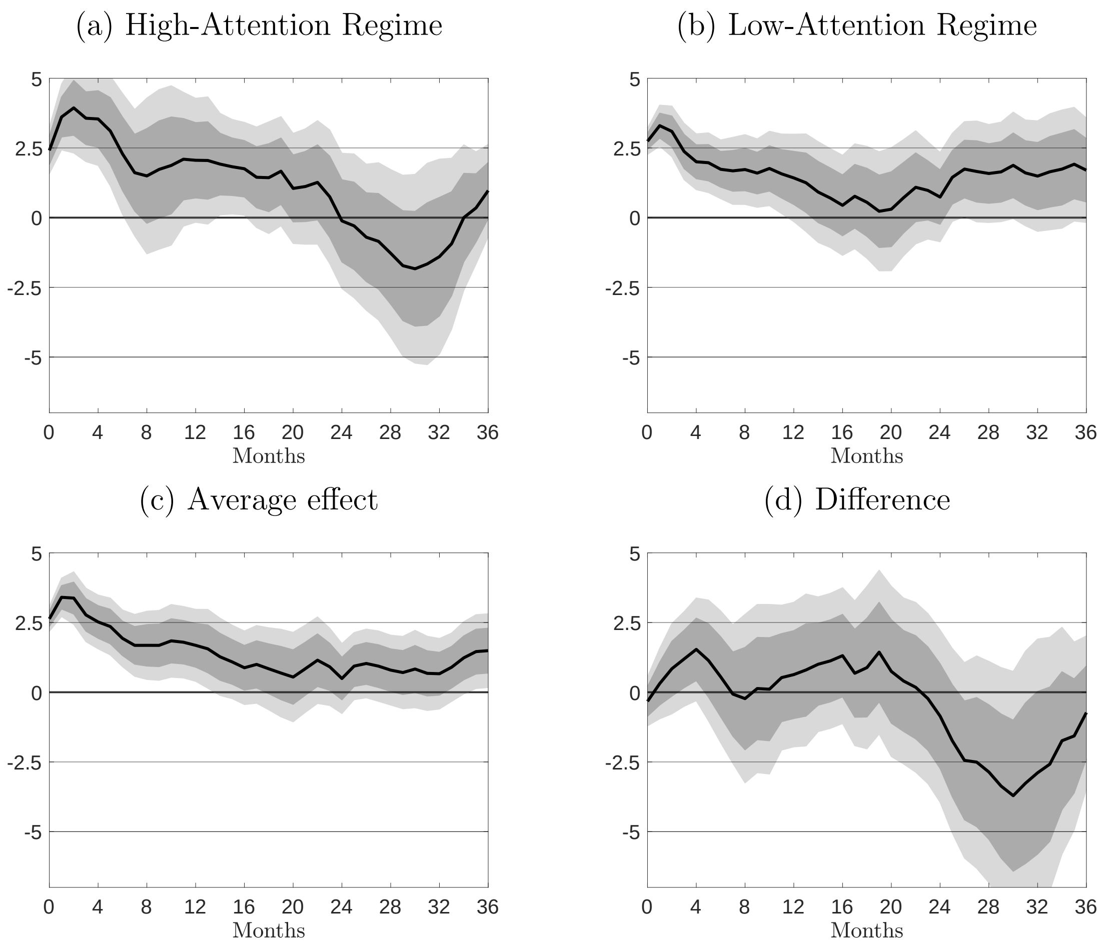


Notes: This figure shows the inflation response to a monetary policy shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags). The shocks are from Jarociński and Karadi (2020) and are based on a high-frequency identification and cleaned from information effects of monetary policy.

C.2 Robustness analysis

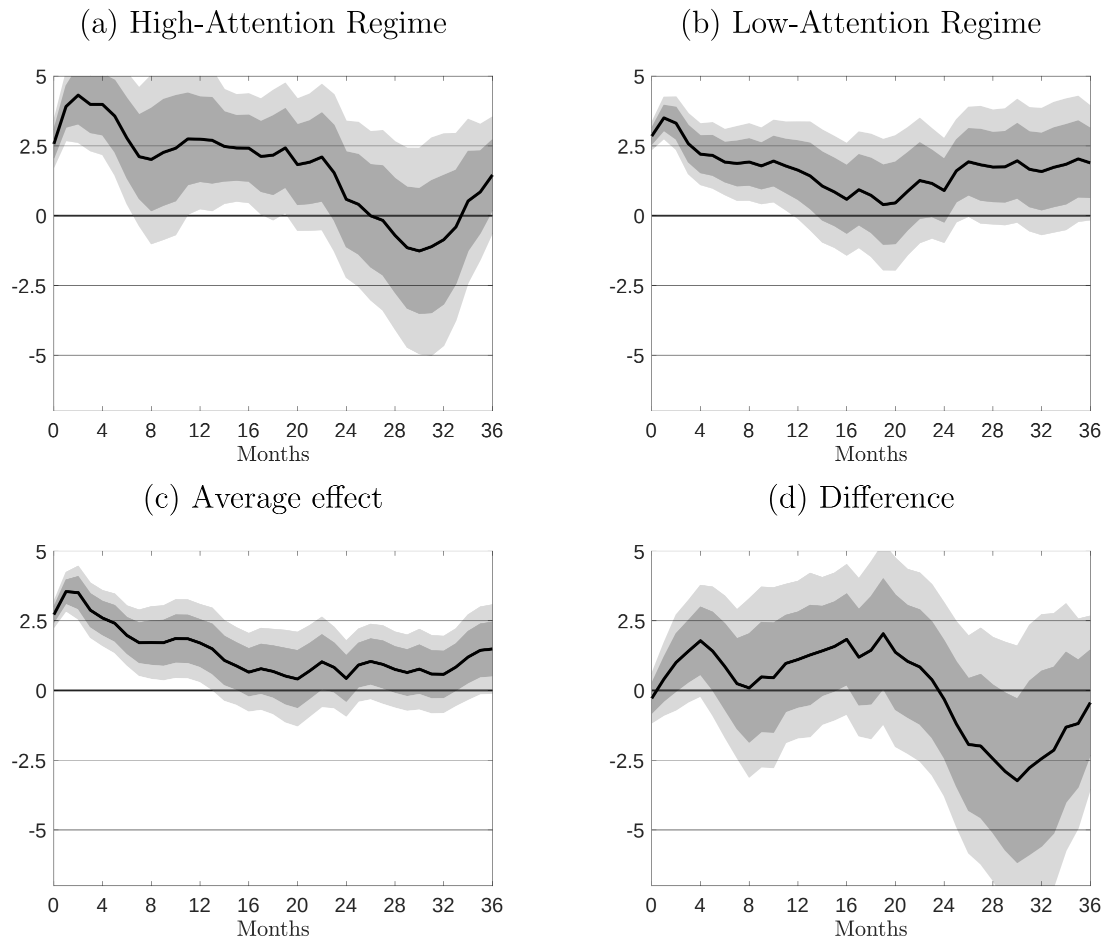
Oil price response. Figures 18 and 19 show the responses of the real and nominal oil price, respectively, to a negative oil supply news shock for the two attention regimes (panels (a) and (b)), the average response in panel (c), and the difference across regimes in (d). We see that the differences across regimes are not significant at the 10% significance level at any horizon, indicating that the differences in the inflation responses discussed in section 3 are unlikely to be driven by different oil price responses.

Figure 18: Real oil price response to an oil supply news shock



Notes: This figure shows the real oil price response to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags). The attention regimes ares defined based on the previous month's inflation rate.

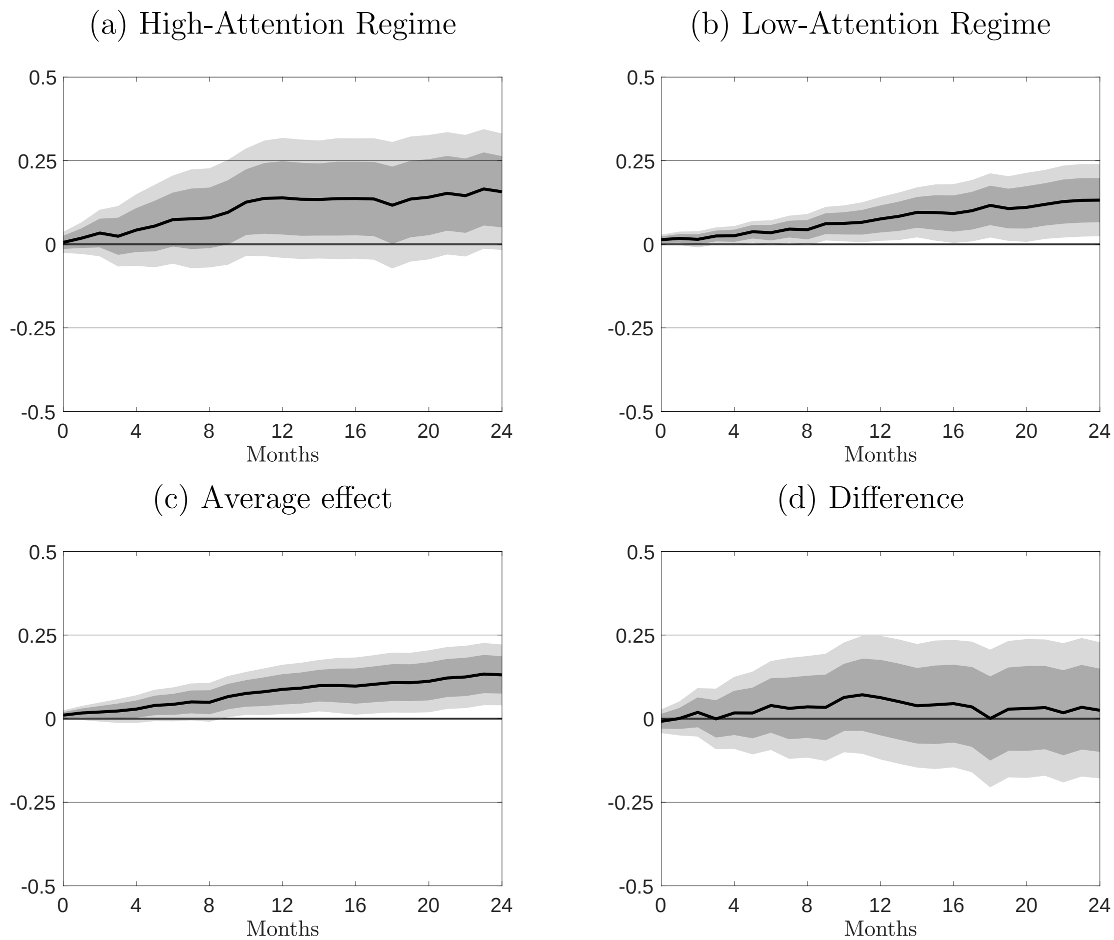
Figure 19: Nominal oil price response to an oil supply news shock



Notes: This figure shows the nominal oil price response to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags). The attention regimes ares defined based on the previous month's inflation rate.

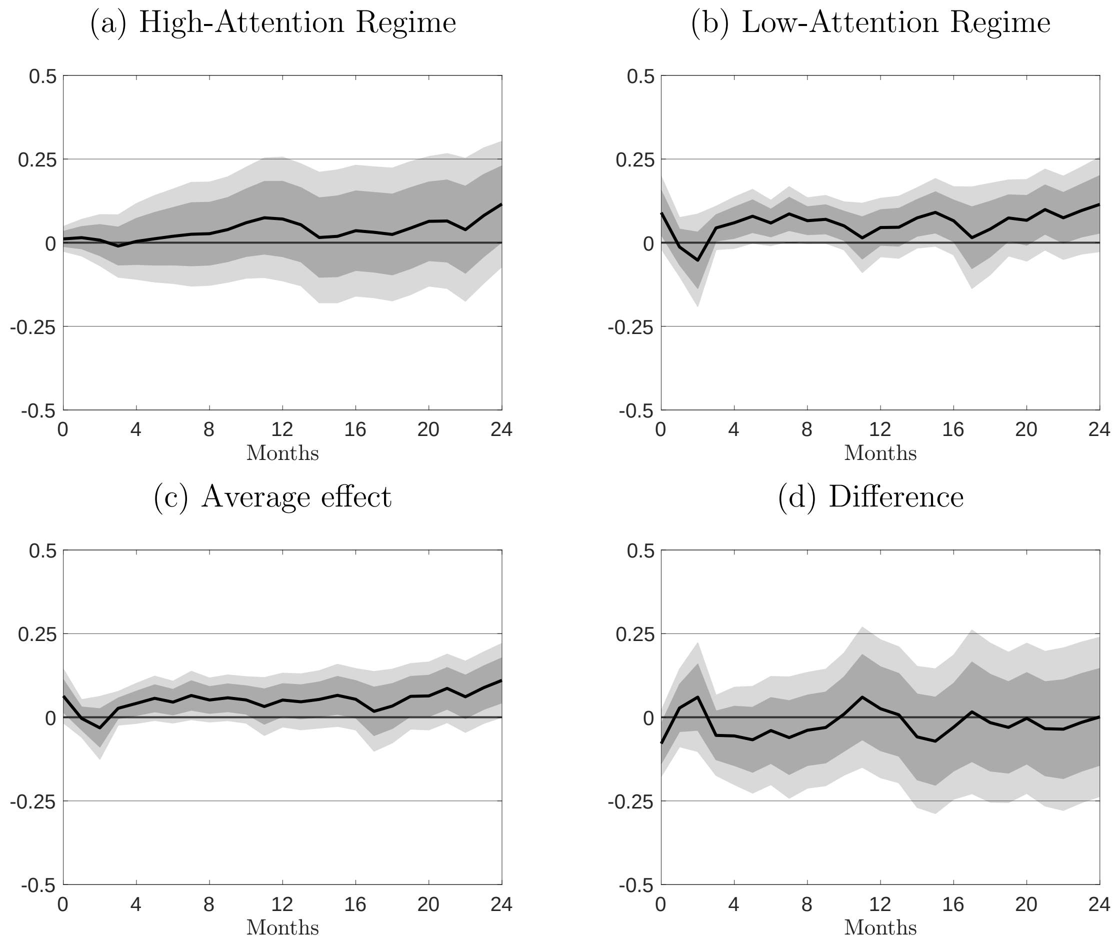
Unemployment response. Figures 20 and 21 show the responses of the unemployment rate (excluding the Covid period 2020-2023 and including it, respectively) to a negative oil supply news shock for the two attention regimes (panels (a) and (b)), the average response in panel (c), and the difference across regimes in (d). We see that the differences across regimes are not significant at any conventional significance level at any horizon, indicating that the differences in the inflation responses discussed in section 3 are unlikely to be driven by generally more responsive macroeconomic variables.

Figure 20: Unemployment response to an oil supply news shock



Notes: This figure shows the unemployment rate response to an oil supply news shock (equation (6)) when excluding the Covid-19 period (2020-2023) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

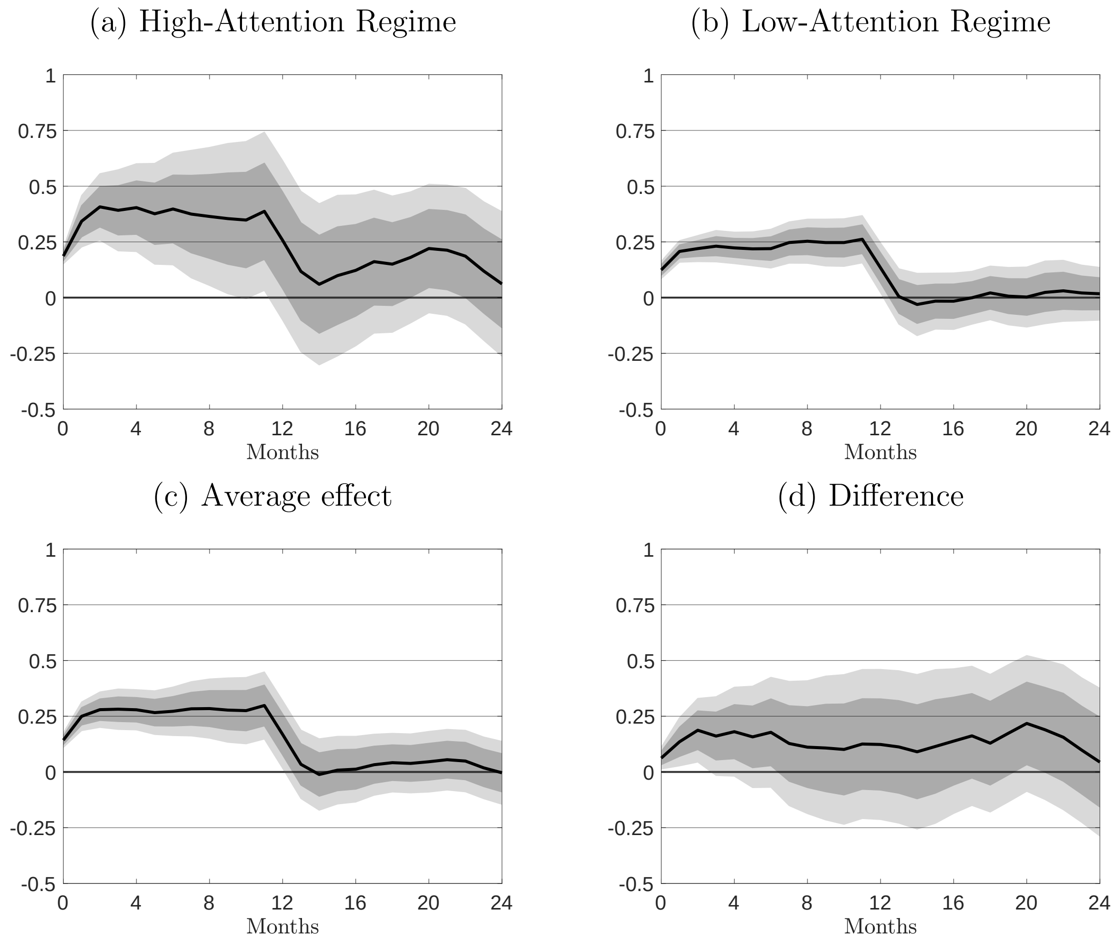
Figure 21: Unemployment response to an oil supply news shock including Covid



Notes: This figure shows the unemployment rate response to an oil supply news shock (equation (6)) when including the Covid-19 period (2020-2023) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Covid. Figure 22 shows the inflation response to a negative oil supply news shock for the two attention regimes (panels (a) and (b)), the average response in panel (c), and the difference across regimes in (d) when excluding the Covid period 2020-2023. We see that the differences across regimes are slightly weaker than for the baseline case in Figure 2 but still substantial.

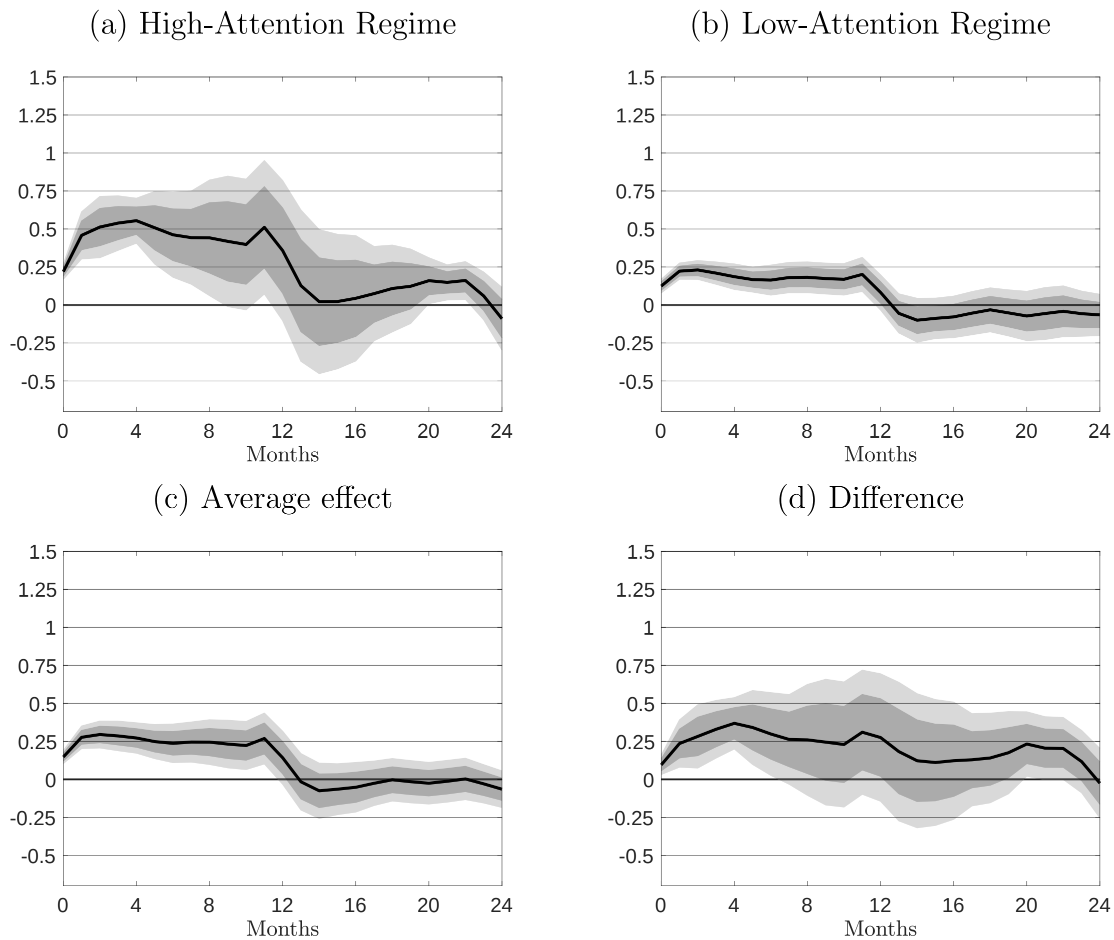
Figure 22: Inflation response to an oil supply news shock excluding Covid



Notes: This figure shows the inflation response to an oil supply news shock (equation (6)) when excluding the years 2020, 2021, 2022 and 2023 in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Exluding the Great Inflation period. Figure 23 shows the inflation response to a negative oil supply news shock for the two attention regimes (panels (a) and (b)), the average response in panel (c), and the difference across regimes in (d) when excluding the Great inflation period, i.e., when excluding observations before 1990. We see that the differences across regimes are even slightly stronger than in the baseline specification.

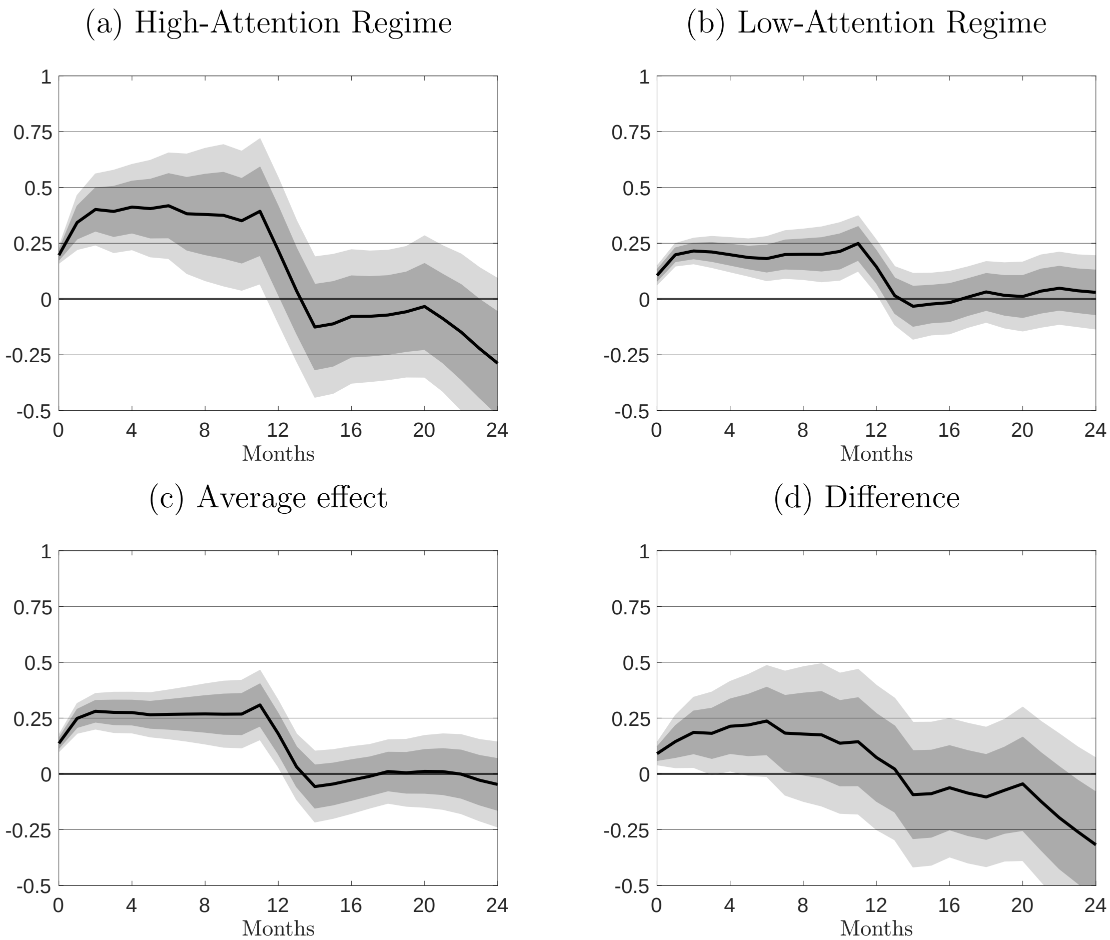
Figure 23: Inflation response to an oil supply news shock after 1990



Notes: This figure shows the inflation response to an oil supply news shock (equation (6)) when excluding the years 2020, 2021, 2022 and 2023 in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Role of control variables. Figure 24 shows the inflation response to a negative oil supply news shock for the two attention regimes (panels (a) and (b)), the average response in panel (c), and the difference across regimes in (d) when not including any control variables. We see that the differences across regimes are slightly weaker than for the baseline case in Figure 2 but still substantial.

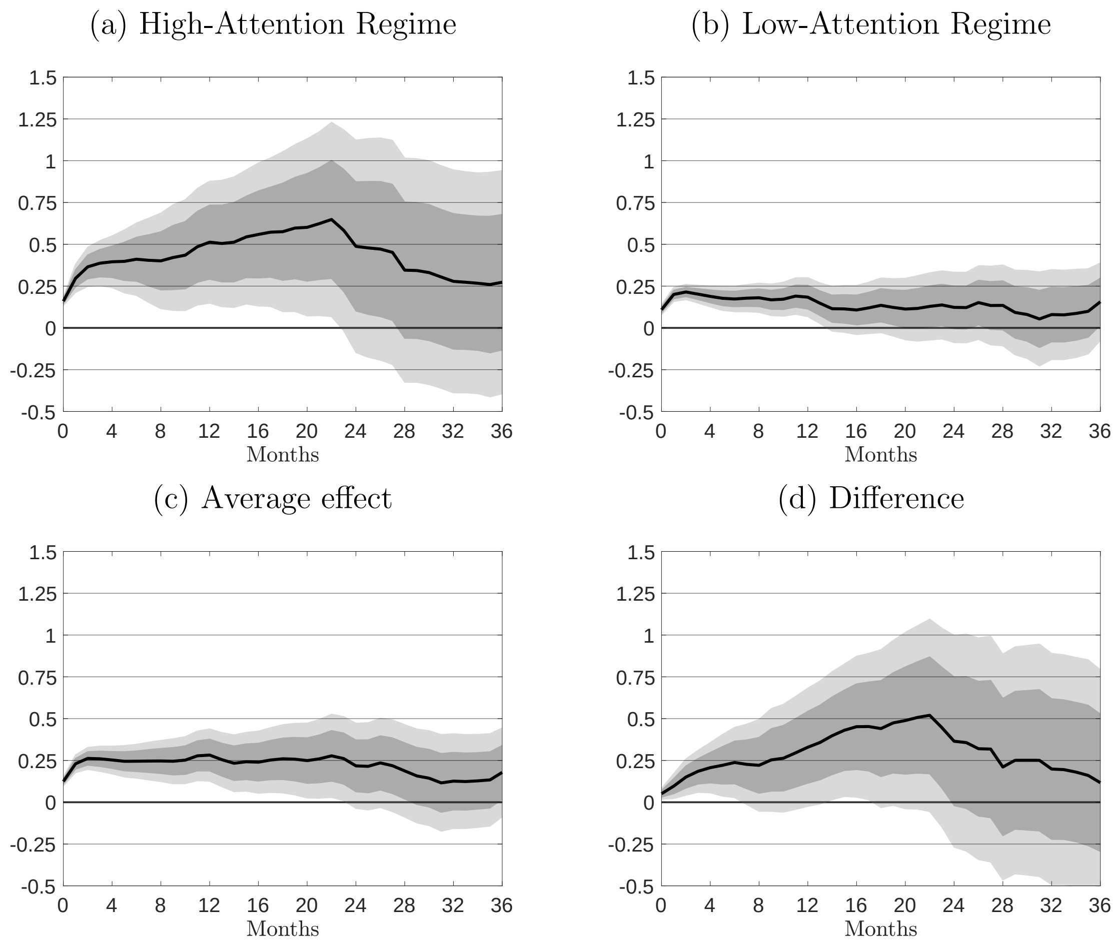
Figure 24: Inflation response to an oil supply news shock (no controls)



Notes: This figure shows the inflation response to an oil supply news shock (equation (6)) when using no control variables in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

Price-level response. In the baseline regression (6), the dependent variable is the change in the inflation rate from time $t-1$ to time $t+h$: $\pi_{t+h} - \pi_{t-1}$. Figure 25 now shows the case where the dependent variable is the cumulative change in the price level: $P_{t+h} - P_{t-1}$, where P_t is the natural logarithm (times 100) of the price level using the CPI. We see from Figure 25 that the results are similar to the baseline case in figure 2: inflation responds 2-3 times more strongly in the high-attention regime and the differences are highly statistically significant, at least in the first 1-2 years.

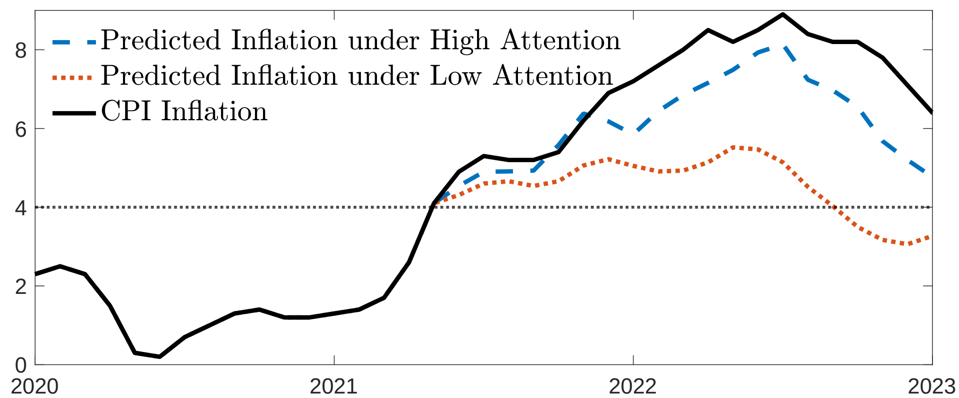
Figure 25: Price level response to an oil supply news shock



Notes: This figure shows the cumulative price level (using the CPI) response to an oil supply news shock (equation (6)) in the high-attention regime (panel (a)), the low-attention regime (panel (c)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). The dark shaded areas depict the 68% confidence bands and the light-shaded area the 90% confidence bands. Standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags). The attention regimes ares defined based on the previous month's inflation rate.

C.3 The role of attention in the recent inflation surge: using Google Trends

Figure 26: Supply shocks, attention and the post-Covid inflation surge



Notes: This figure shows the actual US CPI inflation from 2020 until 2023, as well as the counterfactual inflation dynamics arising only from oil supply news shocks in the economy accounting for the increase in attention in April 2021 (blue-dashed line) and in the economy ignoring the increase in attention (orange-dotted line). The implied inflation dynamics depend on the estimated impulse responses using Google trends data (Figure 3).

D Model Derivations

Derivation of equation (23). Linearizing the Euler equation (9) yields

$$\widehat{c}_t = \tilde{E}_t \widehat{c}_{t+1} - \varphi \left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1}) \right). \tag{34}$$

The representative household therefore needs to form expectations about future consumption $\tilde{E}_t \widehat{c}_{t+1}$, inflation $\tilde{E}_t \pi_{t+1}$ and the exogenous preference shock $\tilde{E}_t \widehat{z}_{t+1}$. The household holds rational expectations about the shock, so $\tilde{E}_t \widehat{z}_{t+1} = E_t \widehat{z}_{t+1}$, where $E_t(\cdot)$ denotes the rational expectations operator, and since $\widehat{z}_t$ follows an AR(1) process with persistence ρ_z, it follows that $\tilde{E}_t \widehat{z}_{t+1} = E_t \widehat{z}_{t+1} = \rho_z \widehat{z}_t$.

Following the discussion in Section 2, the household holds a subjective model of how consumption and inflation evolve. She believes that they both follow a random walk:

$$\pi_t = \pi_{t-1} + \nu_{\pi,t}$$
$$\widehat{c}_t = \widehat{c}_{t-1} + \nu_{c,t},$$

where $\nu_{\pi,t}$ and $\nu_{c,t}$ are normally distributed with mean zero, time-invariant standard deviations and independent from each other. At the time the household needs to form expectations about future consumption and inflation, she does not perfectly observe their current realizations.

This assumption could capture that the household is not at all times perfectly monitoring all consumption expenditures of the members of the household. Instead of observing consumption (and inflation) perfectly, the household only receives noisy signals of the form

$$s_{\pi,t} = \pi_t + \varepsilon_{\pi,t}$$
$$s_{c,t} = \widehat{c}_t + \varepsilon_{c,t},$$

with normally distributed noise terms $\varepsilon_{\pi,t}$ and $\varepsilon_{c,t}$. As explained in Sections 2 and A, the noise of these signals is endogenous and depends on how attentive the household is. Given these signals, it follows from the (steady state) Kalman filter that optimal updating in regime r is given by

$$\tilde{E}_t\pi_{t+1} = \tilde{E}_{t-1}\pi_t + \gamma_{\pi,r}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_{\pi,t}$$
$$\tilde{E}_t\widehat{c}_{t+1} = \tilde{E}_{t-1}\widehat{c}_t + \gamma_{c,r}\left(\widehat{c}_t - \tilde{E}_{t-1}\widehat{c}_t\right) + u_{c,t}.$$

I abstract from noise shocks and therefore impose that $u_{\pi,t}$ and $u_{c,t}$ are zero.

In order to express equation (34) in terms of the output gap, we need to derive the efficient allocation, i.e., the one that prevails in the economy with fully flexible prices. From the production function, we have $Y_t^* = H_t^*$. The real wage is constant $w_t^* = 1$. From the labor-leisure equation (10), we get that potential output is therefore also constant and equal to

$$Y_t^* = \Xi^{-\frac{1}{\nu+\sigma}}. \tag{35}$$

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 directly that

$$r_t^* = \widehat{z}_t - \tilde{E}_t z_{t+1}. \tag{36}$$

Substituting $\widehat{z}_t - \tilde{E}_t z_{t+1}$ with r_t^* in (34) and using that $\widehat{c}_t = \widehat{y}_t = \widehat{x}_t$, since potential output is 0, yields the IS equation (23):

$$\widehat{x}_t = \tilde{E}_t\widehat{c}_{t+1} - \varphi\left(\tilde{i}_t - \tilde{E}_t\pi_{t+1} - r_t^*\right). \tag{37}$$

Note, however, that if we assume initial values $\tilde{E}_{-1}\widehat{c}_0 = \tilde{E}_{-1}\widehat{x}_0$, it follows that $\tilde{E}_t\widehat{c}_{t+1} = \tilde{E}_t\widehat{x}_{t+1}$ because $\widehat{c}_t = \widehat{x}_t$ in equilibrium. This holds true, even if the household does not know that they are equal in equilibrium. We can thus write equation (37) as

$$\widehat{x}_t = \tilde{E}_t \widehat{x}_{t+1} - \varphi \left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - r_t^* \right), \tag{38}$$

which is equation (23) as stated in the main text.

Derivation of equation (22). To derive the New Keynesian Phillips Curve (22), we start by maximizing

$$\Omega_0(j) = \tilde{E}_0 \sum_{t=0}^{\infty} \beta^t \left[(1-\tau_t) P_t(j) \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} \frac{Y_t}{P_t} - w_t H_t(j) - \frac{\psi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right)^2 Y_t + t_t^F(j) \right],$$

which yields the first-order condition

$$\begin{aligned}(1-\tau_t)(\epsilon-1) P_t(j)^{-\epsilon} \frac{Y_t}{P_t^{1-\epsilon}} = \epsilon mc_t \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon-1} \frac{Y_t}{P_t} - \psi \left(\frac{P_t(j)}{P_{t-1}(j)} - 1 \right) \frac{Y_t}{P_{t-1}(j)} \\ + \beta \psi \tilde{E}_t \left[\left(\frac{P_{t+1}(j)}{P_t(j)} - 1 \right) \frac{P_{t+1}(j)}{P_t(j)} \frac{Y_{t+1}}{P_t(j)} \right].\end{aligned}$$

I define $T_t \equiv 1 - \tau_t$, and set it such that marginal costs in the steady state are zero, i.e., $T = \frac{\epsilon}{\epsilon - 1}$. Linearizing around the zero-inflation steady state in which $P(j) = P$ for all j, I obtain

$$T(\epsilon-1)\frac{Y}{P} \left[\widehat{T}_t - \epsilon \widehat{p}_t(j) + \widehat{y}_t - (1-\epsilon) \widehat{p}_t \right] \tag{39}$$

$$= \epsilon mc \frac{Y}{P} \left[(-\epsilon - 1) \widehat{p}_t(j) - (-\epsilon - 1) \widehat{p}_t + \widehat{mc}_t + \widehat{y}_t - \widehat{p}_t \right] \tag{40}$$

$$- \psi \frac{Y}{P} \left(\widehat{y}_t - \widehat{p}_{t-1}(j) - \pi_t(j) - \widehat{y}_t + \widehat{p}_{t-1}(j) \right) \tag{41}$$

$$+ \beta \psi \frac{Y}{P} \tilde{E}_t^j \pi_{t+1}^j. \tag{42}$$

Grouping terms, using $mc = 1$, $T = \frac{\epsilon}{\epsilon - 1}$ and dividing by $\frac{Y}{P}$ yields

$$\widehat{p}_t(j) = \frac{1}{\psi + \epsilon} \left[\psi \widehat{p}_{t-1} + \epsilon \left(\widehat{mc}_t - \widehat{T}_t + \widehat{p}_t \right) + \beta \psi \tilde{E}_t^j \pi_{t+1}^j \right]. \tag{43}$$

Given the assumptions in Section 4, it follows that $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t \pi_{t+1}$ such that

$$\widehat{p}_t(j) = \frac{1}{\psi + \epsilon} \left[\psi \widehat{p}_{t-1} + \epsilon \left(\widehat{mc}_t - \widehat{T}_t + \widehat{p}_t \right) + \beta \psi \tilde{E}_t \pi_{t+1} \right]. \tag{44}$$

Therefore, the optimal price $\widehat{p}_t(j)$ is the same for all firms j, $\widehat{p}_t(j) = \widehat{p}_t$, such that we get

$$\psi \underbrace{\left(\widehat{p}_t - \widehat{p}_{t-1}\right)}_{=\pi_t} = \epsilon \left(\widehat{mc}_t - \widehat{T}_t\right) + \beta\psi\tilde{E}_t\pi_{t+1}. \tag{45}$$

Dividing by ψ yields

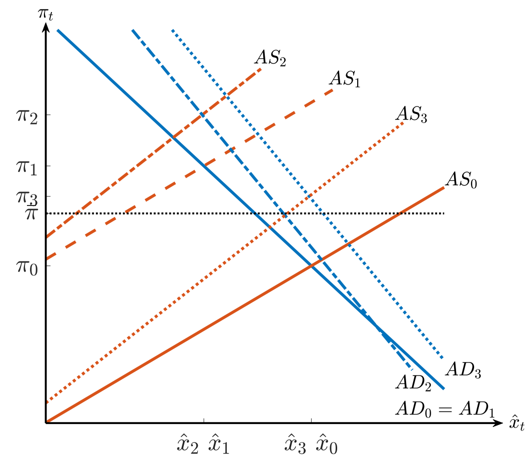
$$\pi_t = \frac{\epsilon}{\psi}\left(\widehat{mc}_t - \widehat{T}_t\right) + \beta\tilde{E}_t\pi_{t+1}. \tag{46}$$

Using that $\widehat{mc}_t = (\sigma + \nu)\,\widehat{y}_t$ and $\widehat{y}_t = \widehat{x}_t$ and defining cost-push shocks as $u_t \equiv -\frac{\epsilon}{\psi}\widehat{T}_t$ and the slope parameter $\kappa \equiv \frac{\epsilon}{\psi}(\sigma + \nu)$, I arrive at the linearized New Keynesian Phillips Curve under subjective expectations:

$$\pi_t = \kappa\widehat{x}_t + \beta\tilde{E}_t\pi_{t+1} + u_t. \tag{47}$$

E Additional Figures

Figure 27: Illustrative example: after the shock has died out



Notes: This figure depicts the third period in which the cost-push shock has died out (after being positive for two consecutive periods). The black-dotted line at $\bar{\pi}$ depicts the inflation attention threshold.

Figure 28: Inflation dynamics after a demand shock

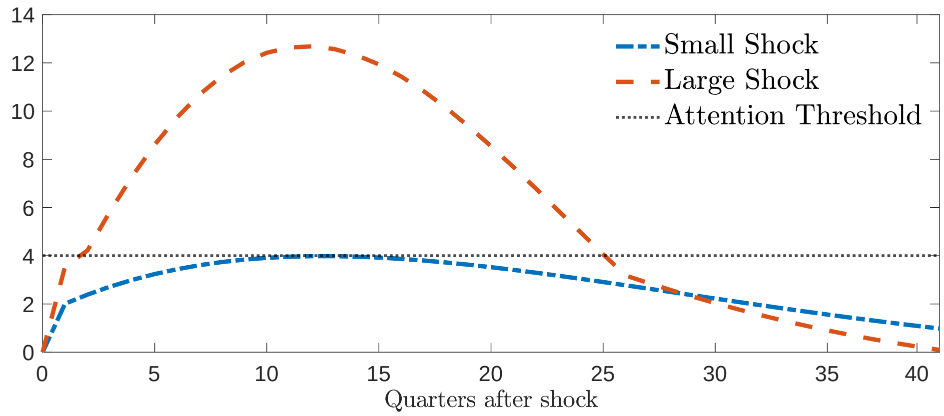


Figure 29: Inflation dynamics under FIRE or absent the attention change

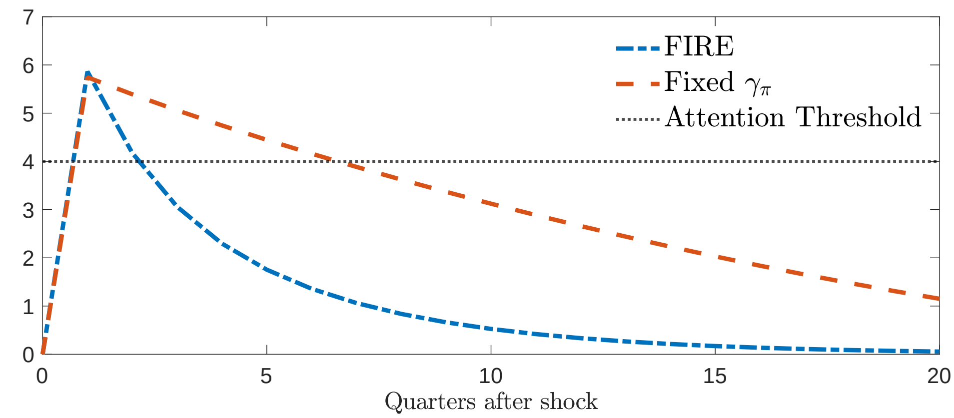


Figure 30: Inflation dynamics with $\gamma_{x,L} = 0.25$ and $\gamma_{x,H} = 0.1$

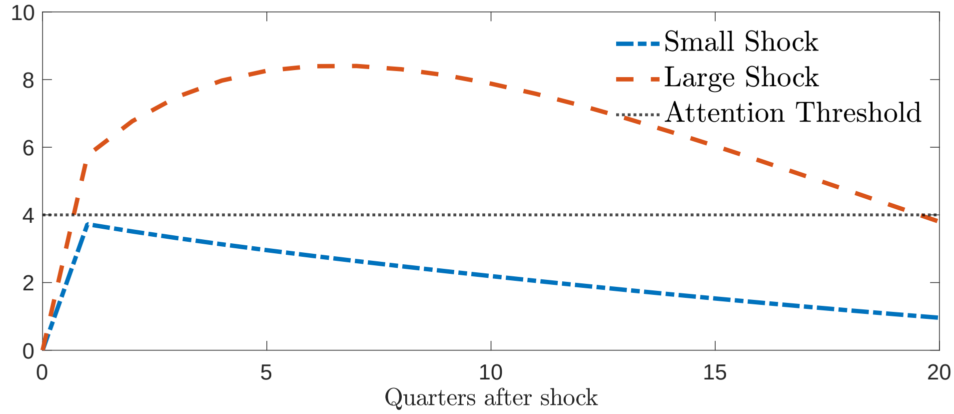


Figure 31: Inflation dynamics with current inflation as the threshold-defining variable

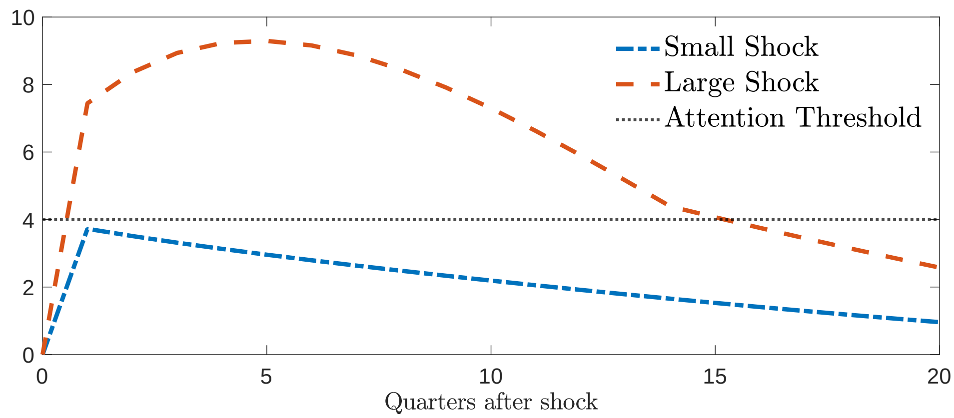


Figure 32: Forecast errors with current inflation as the threshold-defining variable