The Inflation Attention Threshold and Inflation Surges

Oliver Pfäuti*

The University of Texas at Austin

August 2024

Link to most recent version

Abstract

At the outbreak of the recent inflation surge, the public's attention to inflation was low but increased quickly once inflation started to rise. In this paper, I quantify when and by how much the public's attention to inflation changes, and derive the macroeconomic implications of these attention changes. I estimate an attention threshold at an inflation rate of about 4%, and that attention doubles when inflation exceeds this threshold. Adverse supply shocks become more inflationary in times of high attention, and the increase in people's attention to inflation in 2021 accounts for half of the subsequent supply-driven inflation. I develop a model accounting for the attention threshold and show that shocks that are usually short lived lead to a persistent surge in inflation if they induce an increase in people's attention. The attention threshold further lengthens the last mile of disinflation after an inflation surge, and leads to an asymmetry in the dynamics of inflation.

JEL Codes: E3, E4, E5, E7

Keywords: Inattention, Inflation, Inflation Expectations, Monetary Policy

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 (panel (a) in Figure 1), and as inflation peaked, households' and firms' inflation expectations also started to pick up (panel (b) in Figure 1). During the build up of the inflation surge, however, inflation expectations were substantially below actual inflation. Only as inflation started to come down, inflation expectations started to overshoot inflation and these expectations remained high while inflation stayed persistently above its target of 2%.

Figure 1: Inflation, attention to inflation, and inflation expectations

(a) Inflation and Google Trends   (b) Inflation and inflation expectations
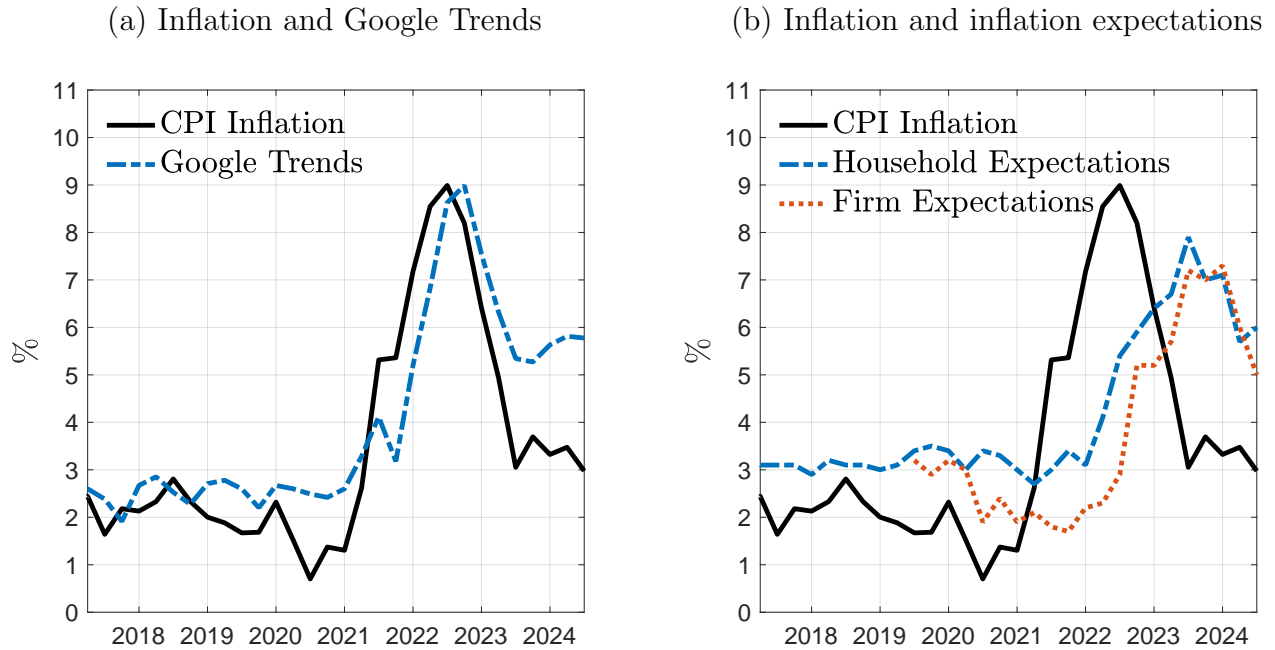
Notes: Panel (a) shows U.S. CPI inflation (black-solid line) together with Google Trends searches of the word 'inflation' in the U.S. Google Trends are normalized to have the same maximum as CPI inflation. Panel (b) shows U.S. CPI inflation (black-solid line) together with one-year-ahead inflation expectations for U.S. households (blue-dashed line) and firms (red-dotted line). Expectations are shifted such that the vertical distance to the black line captures the forecast errors. Household expectations are from the Survey of Consumers from the University of Michigan and firm expectations from the Survey of Firms' Inflation Expectations, provided by the Cleveland Fed.

In this paper, I show that accounting for the observed changes in people's attention to inflation provides an explanation for the recent inflation surge as well as the interplay between inflation and inflation expectations observed in panel (b) of Figure 1. When attention to inflation increases when inflation exceeds a certain threshold, inflation expectations become more sensitive to inflation in times of high inflation and thus, reinforce the inflation surge and render it more persistent. As inflation falls back below the threshold, however, attention decreases and inflation expectations become more

---

[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).

sluggish, leading to a persistent overshooting of inflation expectations and a longer 'last mile' of disinflation back to target.

To arrive at these results, I propose a way to estimate attention to inflation and 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 short-run inflation expectations in response to forecast errors.[2] In this framework, agents may become more attentive because their prior uncertainty increases, because the cost of being attentive decreases, or because the perceived persistence of inflation increases. I provide evidence for all three channels.

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 2024. I find an inflation attention threshold at an annualized inflation rate of about 4%, and that attention is about 0.18 when inflation is below the threshold and practically doubles to 0.35 when inflation is above the threshold. An attention level of 0.35 means that following a 1pp forecast error, agents increase their inflation expectations by 0.35pp. A Markov-switching model, in which I do not need to assume what determines the threshold, confirms these findings. Using regional variation in inflation and inflation expectations, or when relying on the panel data from the Survey of Consumer Expectations from the New York Fed for the period after 2013, I obtain very similar results. I find that the data suggests that there is only one attention threshold, that the threshold specification is preferred to smooth changes in attention, and that once I control for this threshold, I do not find any evidence that attention increases with inflation *within* regime.

The attention regimes matter for the dynamics of inflation and inflation expectations. Using the high-frequency identification of Känzig (2021), I use oil supply news 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 20 basis points on average across regimes, leads to a peak increase in inflation of about 40 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. Using regional data and Google Trends data as a proxy for people's attention, I further disentangle the effects of attention from other changes in the economy that may arise from higher inflation, and find that attention indeed leads to a stronger inflation response to supply shocks even when controlling for inflation directly.

The attention threshold I estimate implies that the U.S. economy entered the high-attention regime most recently in early 2021. To quantify the importance of this increased attention for the recent inflation surge, I first show that the oil supply news shocks explain about 55% of the inflation surge from the time the U.S. 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

---

[2]For example, Werning (2022) discusses why short-run expectations are more important than long-run expectations for inflation. In Pfäuti (forthcoming), I show that measuring attention to inflation by estimating how agents update their short-run expectations correlates strongly with other measures of attention.

have only been half as strong.[3] 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 and rendered inflation more persistent.

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. Attention changes occur because the shock volatility increases once inflation exceeds the threshold. These regime-dependent shock volatilities are calibrated to match the empirically-estimated shock volatility and the increase in attention across regimes jointly. I find that in response to an inflationary supply shock that pushes inflation above the attention threshold, inflation keeps on increasing through the 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, leading to even higher inflation expectations. This illustrates how what normally 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. The inflation attention threshold and changing degrees of attention hence offer an explanation for why 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, as documented in panel (b) of Figure 1. Therefore, the changes in attention to inflation can explain why inflation came down relatively quickly during the second half of 2022 but remained stubbornly above the Fed's inflation target even in early 2024.

That inflation expectations stay persistently high, even when inflation has already fallen back below the threshold, further 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 deflation largely unchanged. This inflation asymmetry induced by the attention threshold also offers a potential explanation for why we did not

---

[3]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. In the model, however, I show that demand shocks can lead to similar inflation dynamics as supply shocks, consistent with my empirical findings.

observe long lasting deflationary periods, for example, during the Great financial crisis (Coibion and Gorodnichenko, 2015b). 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 about 30% of the time during the same period.

This asymmetry also matters for the normative implications of the attention threshold. When monetary policy sets the nominal interest rate following a relatively dovish Taylor rule, the associated central bank losses are significantly larger compared to optimal policy rules or a strict inflation targeting rule. The reason is that under a standard Taylor rule, the economy spends a substantial amount of time in the high-attention regime in which inflation is high and volatile. Being frequently in the high-attention regime also increases the *average level of inflation* due to the asymmetry of the threshold which is costly for the central bank. By following a more hawkish monetary policy rule, the central bank can mitigate these losses.

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. Cerrato and Gitti (2022) and Benigno and Eggertsson (2023) find evidence for a non-linear Phillips Curve that is steeper when labor markets are tight. Relatedly, Ball et al. (forthcoming) find that the tightening of labor markets was a main driver of the rise in core inflation, and Lorenzoni and Werning (2023) focus on wage-price spirals. Crump et al. (2024) document a prominent role of expected unemployment gaps for the observed inflation dynamics, Erceg et al. (2024) focus on non-linear wage and price Phillips curves coupled with indexation and learning about the persistence of the shocks. 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. Cavallo et al. (forthcoming) and Blanco et al. (2024) argue that the frequency of price adjustments has increased. In contrast to all of these papers, I focus on the role of changes in people's attention to inflation. Attention serves as a propagation mechanism rather than as an exogenous impulse and my results show that the increase in attention in early 2021 roughly doubled the inflationary effects of supply shocks between early 2021 and the end of 2022.[4] Additionally, my findings offer an explanation for the interplay between inflation and inflation expectations (panel (b) in Figure 1) as well as for the long last mile of disinflation.

---

[4]See Binder and Kamdar (2022) for a narrative history of how inflation and inflation expectations evolved jointly during several recessions and periods of high inflation since the Great Depression.

Pfäuti (forthcoming) documents that before the recent inflation surge, attention to inflation was at a historical low.[5] Bracha and Tang (forthcoming) find that when inflation increases, attention to inflation increases as well. Korenok et al. (forthcoming) 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 U.S. Buelens (2024) uses internet search behavior of households in the Euro Area to show that this form of attention to inflation positively depends on the level of inflation. 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 (2024) 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 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 macroeconomic implications of these attention changes.

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; Aastveit et al., 2023; Jo and Zubairy, forthcoming). 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.[6] 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. Turen (2023) shows how a model with

---

[5]For the years before the recent inflation surge, Candia et al. (forthcoming, 2023) and Coibion et al. (2023) 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 Coibion and Gorodnichenko (2015a) who show state-dependence in the degree of information rigidity (focusing on GDP growth), or 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). York (2023) exploits differences in survey dates to examine how inflation expectations respond to macroeconomic news releases.

[6]Based on the present paper and Korenok et al. (forthcoming), 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.

state-dependent attention can reconcile several facts about the microeconomic evidence on price setting.[7] Gáti (2023) 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 (forthcoming), 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 (forthcoming) 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 (2021) 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 estimate and incorporate 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 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 discuss the threshold's normative implications and Section 7 concludes.

2   Attention and the Inflation Attention Threshold

In this section, I derive a model of optimal attention choice which I then bring to the data to estimate the inflation attention threshold and the different degrees of attention in the two attention regimes.

2.1   Attention to inflation

In order to estimate attention to inflation as well as a threshold of when attention changes, I extend the method I develop in Pfäuti (forthcoming) by allowing for state dependency of the model's

---

[7]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, and Lorenzoni (2009) focuses on dispersed information.

parameters.[8] The agent believes that inflation in the next period, π', depends on inflation today, π, as follows

$$\pi' = \rho_{\pi,r}\pi + \nu,$$

where $\nu \sim i.i.N.(0, \sigma_\nu^2)$, and $\rho_{\pi,r} \in [0,1]$ denotes the perceived persistence of inflation which may depend on the state or regime r that the economy is in.[9] Inflation in the current period is unobservable, so before forming an expectation about future inflation, the agent needs to form expectations about today's inflation. I denote this nowcast $\widetilde{\pi}$, and the resulting forecast about next period's inflation $\pi^e = \rho_{\pi,r}\widetilde{\pi}$. The assumption that inflation expectations solely depend on perceived inflation is supported by a large empirical literature documenting that perceived inflation is indeed the key predictor for inflation expectations (see, e.g., Jonung (1981) and Weber et al. (2022) for households and Candia et al. (2023) for firms). That said, however, I also allow for a multivariate expectations formation process in the estimation later. Given these beliefs, the full-information forecast π^{e*} is $\pi^{e*} \equiv \rho_{\pi,r}\pi$. But because π is not perfectly observable, the actual forecast will deviate from this full-information forecast. These deviations are costly and the agent acquires information in order to prevent these deviations. Thus, 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.[10]

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

$$\max_f \int U(\pi^e, \pi)f(\pi^e, \pi)d\pi d\pi^e - C(f) \tag{1}$$

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

where $g(\pi)$ is the agent's prior, which is assumed to be Gaussian; $\pi \sim N\left(\widehat{\pi}, \sigma_{\pi,r}^2\right)$, where I also allow prior uncertainty $\sigma_{\pi,r}^2$ to be state or regime dependent. $C(.)$ 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:

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

---

[8]The method proposed in Pfäuti (forthcoming) builds on Mackowiak et al. (2023) and Vellekoop and Wiederholt (2019), see also Weber et al. (2023) for a similar model but with the goal to explain their RCT findings.

[9]Agents' subjective model does not necessarily need to 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).

[10]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.

where $H(x) = -\int f(x)log(f(x))dx$ is the entropy of x and $\frac{1}{\lambda_r}$ measures the cost of information, which may also depend on the current state.

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

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

where χ measures the stakes of making a mistake.

In this setup, Gaussian signals are optimal (and in fact the unique solution, see Matějka and McKay (2015)) and take the form

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

with $\varepsilon \sim i.i.N.(0, \sigma_{\varepsilon,r}^2)$. The problem (1) now reads

$$\max_{\sigma_{\pi|s}^2 \leq \sigma_{\pi,r}^2} E_\pi \left[E_s \left[-\chi\rho_{\pi,r}^2 \left(\pi - E[\pi|s]\right)^2\right]\right] - \frac{1}{\lambda_r}I(\pi; \pi^e) = \max_{\sigma_{\pi|s}^2 \leq \sigma_{\pi,r}^2} \left(-\chi\rho_{\pi,r}^2\sigma_{\pi|s}^2 - \frac{\frac{1}{\lambda_r}}{2}log\frac{\sigma_{\pi,r}^2}{\sigma_{\pi|s}^2}\right). \quad (2)$$

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

$$\pi^e = \rho_{\pi,r}\left(1 - \gamma_{\pi,r}\right)\widehat{\pi}_r + \rho_{\pi,r}\gamma_{\pi,r}s, \quad (3)$$

where $\gamma_{\pi,r} = 1 - \frac{\sigma_{\pi|s}^2}{\sigma_{\pi,r}^2} \in [0,1]$ measures the agent's attention to inflation in state or regime r, and $\widehat{\pi}_r$ denotes the prior mean of π. When the agent is inattentive, she obtains relatively noisy signals and thus, puts little weight on these signals, reflected in a small $\gamma_{\pi,r}$. Therefore, lower attention, i.e., a lower $\gamma_{\pi,r}$, implies that the agent updates her expectations to a given signal s less strongly.

Now, since the agent chooses the level of attention, we can re-formulate (2) as

$$\max_{\gamma_{\pi,r}\in[0,1]} \left(-\chi\rho_{\pi,r}^2(1 - \gamma_{\pi,r})\sigma_{\pi,r}^2 - \frac{\frac{1}{\lambda_r}}{2}log\frac{1}{1 - \gamma_{\pi,r}}\right), \quad (4)$$

which yields the *optimal* level of attention

$$\gamma_{\pi,r} = max\left\{0, 1 - \frac{\frac{1}{\lambda_r}}{2\chi\rho_{\pi,r}^2\sigma_{\pi,r}^2}\right\}. \quad (5)$$

Expression (5) illustrates that attention to inflation depends on the economic environment. I focus on the case in which attention fluctuates between two regimes $r \in \{L, H\}$ depending on the current level of inflation. Later, I will test and confirm that the data indeed prefers this two-regime specification vis-a-vis specifications with multiple thresholds or other forms of changing degrees of attention. Such a regime shift occurs if either (i) the cost of information acquisition and processing decreases as inflation exceeds a certain threshold $\bar{\pi}$, (ii) the perceived persistence increases, or (iii) when $\sigma_{\pi,r}^2$

increases with the threshold:

$$\text{(i) } \frac{1}{\lambda_r} = \begin{cases} \frac{1}{\lambda_L}, & \text{if } \pi < \bar{\pi} \\ \frac{1}{\lambda_H}, & \text{if } \pi \geq \bar{\pi} \end{cases} , \text{ or (ii) } \rho_{\pi,r} = \begin{cases} \rho_{\pi,L}, & \text{if } \pi < \bar{\pi} \\ \rho_{\pi,H}, & \text{if } \pi \geq \bar{\pi} \end{cases} , \text{ or (iii) } \sigma_{\pi,r}^2 = \begin{cases} \sigma_{\pi,L}^2, & \text{if } \pi < \bar{\pi} \\ \sigma_{\pi,H}^2, & \text{if } \pi \geq \bar{\pi}, \end{cases}$$

with at least one of the following holding with inequality: $\frac{1}{\lambda_L} \geq \frac{1}{\lambda_H}$, $\rho_{\pi,L} \leq \rho_{\pi,H}$ and $\sigma_{\pi,L}^2 \leq \sigma_{\pi,H}^2$. Later, I also discuss that all of the three possible drivers of changing degrees of attention are supported in the data, but that given the evidence in Weber et al. (2023), channel (iii) is the most likely one and will be the one I focus on in the New Keynesian model later. The expression for optimal attention (5) has the main purpose of providing testable predictions of what drives attention, and helps guiding the model calibration, but note, that my (empirical and model-implied) results do not depend on the particular analytical expression (5) or the exact driver of the observed attention changes.

Dynamic model. To bring the attention model to the data, I now introduce dynamics. The agent believes that inflation π follows

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

where $\underline{\pi}_r$ is the agent's long-run belief about inflation which may also be regime dependent. 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 normally distributed with variance $\sigma_{\varepsilon,r}^2$. As in the static model, higher attention is reflected in less noise, i.e., a lower $\sigma_{\varepsilon,r}^2$.

I assume that agents always use the steady state Kalman filter in each regime r, so that optimal updating in a given regime r is given by

$$\tilde{E}_t\pi_{t+1} = (1 - \rho_{\pi,r})\underline{\pi}_r + \rho_{\pi,r}\tilde{E}_{t-1}\pi_t + \rho_{\pi,r}\gamma_{\pi,r}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, \tag{6}$$

where $\gamma_{\pi,r}$ captures how attentive the agent is.[11] Explicitly accounting for the attention threshold $\bar{\pi}$, agents form their one-period ahead inflation expectations according to:

$$\tilde{E}_t\pi_{t+1} = \begin{cases} (1 - \rho_{\pi,L})\underline{\pi}_L + \rho_{\pi,L}\tilde{E}_{t-1}\pi_t + \rho_{\pi,L}\gamma_{\pi,L}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, & \text{when } \pi_{t-1} < \bar{\pi} \\ (1 - \rho_{\pi,H})\underline{\pi}_H + \rho_{\pi,H}\tilde{E}_{t-1}\pi_t + \rho_{\pi,H}\gamma_{\pi,H}\left(\pi_t - \tilde{E}_{t-1}\pi_t\right) + u_t, & \text{when } \pi_{t-1} \geq \bar{\pi}. \end{cases} \tag{7}$$

---

[11]Assuming that agents always use the steady state Kalman filter 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); Mackowiak et al. (2023)). This leaves conditional second moments time-invariant and thus, the optimal level of attention constant. An important recent paper that relaxes this assumption is Afrouzi and Yang (2021). However, they focus on the linear-quadratic-Gaussian case which in my case is not applicable due to the non-linearity introduced by the attention threshold.

Note, that I allow the long-run mean belief, the perceived persistence and attention to differ across regimes. Later on, I test for multiple thresholds, changes of attention within regime as well as that attention may depend linearly on the level of inflation or inflation volatility. I find however that the data prefers specification (7). Additionally, I will also allow people's inflation expectations to depend on unemployment and their unemployment expectations. My results are robust to this alternative specification.

2.2 Estimating attention and the attention threshold

Data. As my baseline measure of inflation expectations, I rely on the Survey of Consumers from the University of Michigan. I use average and median household inflation expectations for the period 1978M1-2024M5. For the period 2013-2024, 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 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 transform them into one-quarter-ahead forecasts as follows: $\tilde{E}_t \pi_{t+3} \equiv \frac{\tilde{E}_t \pi_{t+12}}{4}$. Using $\tilde{E}_t \pi_{t+3} \equiv \left[\left(1 + \tilde{E}_t \pi_{t+12} \right)^{1/4} - 1 \right]$ 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: $\pi_t \equiv \frac{P_t - P_{t-3}}{P_{t-3}}$. As I show in Appendix B, the results are qualitatively similar when using one-year-ahead expectations and year-on-year inflation as the actual measure of inflation. I also discuss heterogeneity in attention across demographic groups in Appendix B.

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:

$$
\begin{aligned}
\tilde{E}_t \pi_{t+3} = \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}} \left(\beta_{0,L} + \beta_{1,L} \tilde{E}_{t-3} \pi_t + \beta_{2,L} \left(\pi_t - \tilde{E}_{t-3} \pi_t \right) \right) \\
+ (1 - \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}}) \left(\beta_{0,H} + \beta_{1,H} \tilde{E}_{t-3} \pi_t + \beta_{2,H} \left(\pi_t - \tilde{E}_{t-3} \pi_t \right) \right) + \epsilon_t,
\end{aligned}
\tag{8}
$$

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. I therefore not only allow the attention parameters to differ across regimes, but also the perceived persistence of inflation as well as agents' long-run beliefs about inflation. $\mathbb{1}_{\pi_{t-1} \leq \bar{\pi}}$ 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 $\hat{\bar{\pi}} = 3.91\%$. 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 $\widehat{\gamma}_{\pi,H} = 0.35$. 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.* The threshold value of 3.91% implies that the U.S. economy spent about 30% of the time between 1978 and 2024 in the high-attention regime. Most recently, inflation exceeded this threshold in early 2021. Google searches also started to increase around that time (see panel (a) in Figure 1 in the Introduction).

To test for multiple threshold, I use the Bayesian Information Criterion to select the numbers of thresholds, and I find that the data prefers the specification with only one threshold.[12]

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	3.91%	0.18	0.35	0.000
s.e.		(0.018)	(0.042)	
Median expectations	4.44%	0.13	0.21	0.014
s.e.		(0.018)	(0.027)	
Quarterly frequency	3.21%	0.14	0.38	0.000
s.e.		(0.033)	(0.076)	

Notes: This table shows the results from regression (8), 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. Standard errors are robust with respect to heteroskedasticity.

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. Median expectations tend to be less volatile than average expectations, therefore, it is not surprising that the estimated attention levels—capturing the updating gains—are lower for median expectations. Nevertheless, in both cases, the null hypothesis that the two attention levels across regimes are equal is rejected at most conventional significance levels (with p-values of 0.000 and 0.014, respectively, see last column). Again, I find that the Bayesian information criterion selects having one threshold only.

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 the baseline monthly specification and

---

[12]For example, the BIC increases from -534.12 to -524.70 when going from one to two thresholds and to -512.03 when going to three thresholds. The BIC is given by $T \ln (SSR/T) + k \ln (T)$, where k is the number of parameters, T the number of periods, and SSR the sum of squared residuals. As a robustness check, I also consider the Hannan-Quinn IC and find that also the HQIC selects the model with one threshold only.

again, the difference in attention across regimes is highly statistically significant.

When I use average inflation of the last three months as the threshold-defining variable instead of the last month's inflation rate, I estimate a threshold at 3.45% and attention levels of $\widehat{\gamma}_{\pi,L} = 0.19$ and $\widehat{\gamma}_{\pi,H} = 0.33$.

A drawback of the Survey of Consumers is that it surveys households at most twice. Therefore, it is not possible to examine how individual households update their expectations and how they change their updating behavior with the level of inflation. Focusing on mean (and median) expectations helps to overcome this issue but it also ignores potentially important cross-sectional information. Therefore, I also estimate the attention levels using individual consumer inflation expectations from the Survey of Consumer Expectations. When last month's inflation was below the threshold of 3.91%, I estimate an attention level of 0.08. When inflation was above the threshold, attention increases to 0.32, and I reject the nullhypothesis that the two estimates are equal (p-value of 0.000). Thus, I obtain similar, though slightly lower, attention parameters than in my baseline estimates even though this survey is only available since 2013.

When using the firm expectations from the Survey of Firms' Inflation Expectations (SoFIE) from the Cleveland Fed—with the caveat that that data is only available at quarterly frequency and only since 2018Q2, resulting in only 23 observations—I estimate that attention of firms has increased from practically 0 up to 0.48 as inflation exceeded the threshold of 3.91%.

I also consider mean expectations from the Survey of Professional Forecasters from the Philadelphia Fed. CPI inflation expectations from that survey are available at quarterly frequency since 1981Q3. I estimate an attention threshold at 3.92% and thus, very close to the one for households. The estimated attention levels are $\widehat{\gamma}_{\pi,L} = 0.07$ and $\widehat{\gamma}_{\pi,H} = 0.17$ (with an associated p-value that the two are equal of 0.008).

A multivariate system of inflation expectations. As the perceived law of motion of inflation is an AR(1) process, inflation expectations under limited attention only depend on prior beliefs and inflation itself. In principle, however, it could be possible that inflation expectations also depend on other variables such as unemployment.[13] To take this into account, I therefore now estimate the regression

$$\tilde{E}_t \pi_{t+1} = \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}} \left[\beta_{0,L} + \beta_{1,L} \tilde{E}_{t-1} \pi_t + \beta_{2,L} \left(\pi_t - \tilde{E}_{t-1} \pi_t \right) \right. \tag{9}$$

$$+ \left. (1 - \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}}) \left[\beta_{0,H} + \beta_{1,H} \tilde{E}_{t-1} \pi_t + \beta_{2,H} \left(\pi_t - \tilde{E}_{t-1} \pi_t \right) \right] \right. \tag{10}$$

$$+ \beta_3 \tilde{E}_{t-1} U_t + \beta_4 \left(U_t - \tilde{E}_{t-1} U_t \right) + \tilde{\epsilon}_t, \tag{11}$$

---

[13]Kamdar and Ray (2024), for example, show that consumers tend to expect disinflation when expecting an economic expansion.

where $\tilde{E}_{t-1}U_t$ denote expectations about unemployment changes and U_t are unemployment changes (see Appendix D.1 for further details). I find that both, β_3 and β_4, are insignificant ($\hat{\beta}_3 = -0.13$ and $\hat{\beta}_4 = 0.04$ with p-values of 0.332 and 0.337, respectively). The estimated attention-to-inflation parameters are also barely affected. In fact, I estimate attention in the low regime to be equal to 0.18 and 0.34 in the high regime.

Regional variation. The Survey of Consumers also provides the regions consumers are in, which allows me to leverage regional variation to test for the robustness of my results. In particular, I use the four census regions West, North East, North Central (or Midwest) and South. I control for region fixed effects and estimate that attention is 0.22 with a standard error of 0.023 (standard errors are robust with respect to heteroskedasticity as well as serial and cross-sectional correlation) when last-period's inflation is below the threshold of 3.91%. Attention increases to 0.42 (s.e. of 0.051) when inflation exceeds the threshold. The two estimates are highly statistically significantly different from each other (p-value of 0.000). These results indicate that my baseline results remain robust when considering regional variation.

Markov-switching model. Instead of estimating regression (8), I also estimate a Markov-switching dynamic regression model in which the long-run inflation beliefs $\underline{\pi}$, the perceived persistence ρ_π and attention to inflation γ_π are allowed to differ across unobserved states. That implies that I do not need to take a stand on what defines the regimes, in contrast to equation (8) where I assumed that it is last period's inflation rate that determines the regimes (see Chapter 22 of Hamilton (1994) for a textbook treatment).

The implied attention estimates are 0.19 and 0.33 for the two states and thus, close to the estimates from regression (8) of 0.18 and 0.35. When I allow the variance to change with the state, I estimate attention parameters of 0.2 and 0.32. The implied probability of staying in the low-attention regime from the Markov-switching model is 70% which aligns well with the fact that annualized quarter-on-quarter CPI inflation was below the threshold 70% of the time.

Smooth changes in attention. My baseline regression imposes that attention to inflation takes on only two values and jumps from one to the other when inflation exceeds the threshold or falls back below it. Another possibility is that attention changes smoothly with inflation or inflation volatility. To test for this, I estimate the two alternative regressions:

$$\widetilde{E}_t \pi_{t+3} = \beta_0 + \beta_1 \widetilde{E}_{t-3}\pi_t + \beta_2 \left(\pi_t - \widetilde{E}_{t-3}\pi_t\right) + \beta_3 \pi_{t-1}\left(\pi_t - \widetilde{E}_{t-3}\pi_t\right) + \varepsilon_t \tag{12}$$

$$\widetilde{E}_t \pi_{t+3} = \delta_0 + \delta_1 \widetilde{E}_{t-3}\pi_t + \delta_2 \left(\pi_t - \widetilde{E}_{t-3}\pi_t\right) + \delta_3 \left(\pi_{t-1} - \pi_{t-2}\right)^2 \left(\pi_t - \widetilde{E}_{t-3}\pi_t\right) + \varepsilon_t, \tag{13}$$

where β_3 captures that attention may change with the past level of inflation and δ_3 that it may change with inflation volatility, here measured as the squared change in past inflation.

Again, I use the Bayesian Information Criterion to select which model fits the data best and I find that the data prefers the threshold specification (8) over the two specifications (12) and (13). Additionally, even though the estimated β_3 and δ_3 are statistically significantly different from 0, they are close to 0 (0.009 and -0.002, respectively). In the following, I also test whether a combination of such a 'smooth-attention model' with my threshold model would perform better.

Attention changes within regime. In Table 1, we saw that attention increases when inflation exceeds the threshold of 3.91%. 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 two or five years, respectively.[14] 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

$$\widehat{\gamma}_{\pi,t} = \delta_0 + \delta_1 \mathbb{1}_{\pi_t^{avg} \geq 3.91} + \delta_2 \pi_{t-1} + \delta_3 \mathbb{1}_{\pi_t^{avg} \geq 3.91} \pi_{t-1} + \varepsilon_t, \tag{14}$$

where $\mathbb{1}_{\pi_t^{avg} \geq 3.91}$ is an indicator that equals one when in period t average inflation over the last twelve months is above 3.91% and zero otherwise. Thus, δ_1 tells us the difference of attention across regimes, δ_2 the effect of last-periods inflation on attention and δ_3 the additional effect of inflation on attention in the high-inflation regime.

Table 2 shows the results. We see that attention is significantly higher in the high-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. This holds regardless of whether we use two-year or five-year windows.

Table 2: Attention changes within regime

	$\widehat{\delta}_1$	$\widehat{\delta}_2$	$\widehat{\delta}_3$
2-year windows	0.193***	-0.002	0.022
s.e.	(0.067)	(0.005)	(0.032)
5-year windows	0.119***	-0.004	0.001
s.e.	(0.030)	(0.003)	(0.005)

Notes: This table shows the results from regression (14) when using two- or five-year windows. 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.

An alternative way to test for changes of attention in addition to the jumps across regimes is to

---

[14]To be consistent with the theory, I impose that attention is between 0 and 1, see expression (5). In the case of five-year windows, all estimated attention parameters are between 0 and 1 without having to impose it.

estimate the regression

$$
\begin{aligned}
\tilde{E}_t \pi_{t+1} =& \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}} \left(\beta_{0,L} + \beta_{1,L} \tilde{E}_{t-1} \pi_t + \beta_{2,L} \left(\pi_t - \tilde{E}_{t-1} \pi_t \right) \right) \\
&+ \left(1 - \mathbb{1}_{\pi_{t-1} \leq \bar{\pi}} \right) \left(\beta_{0,H} + \beta_{1,H} \tilde{E}_{t-1} \pi_t + \beta_{2,H} \left(\pi_t - \tilde{E}_{t-1} \pi_t \right) \right) \\
&+ \beta_3 \left(\pi_t - \tilde{E}_{t-1} \pi_t \right) \cdot \pi_{t-1} + \beta_4 \tilde{E}_{t-1} \left[\pi_t \right] \cdot \pi_{t-1} + \tilde{\epsilon}_t,
\end{aligned}
$$

where β_3 and β_4 capture that the perceived persistence and the attention parameter may depend on last-period's inflation in addition to the changes across regimes. I estimate $\hat{\beta}_3 = 0.0037$ (with a standard error of 0.0033 and a p-value of 0.26), and $\hat{\beta}_4 = -0.0031$ (with a standard error of 0.0057 and a p-value of 0.60). When ignoring the effects coming from β_3 and β_4, the attention parameters are $\hat{\gamma}_{\pi,L} = \frac{\hat{\beta}_{2,L}}{\hat{\beta}_{1,L}} = 0.20$ and $\hat{\gamma}_{\pi,H} = \frac{\hat{\beta}_{2,H}}{\hat{\beta}_{1,H}} = 0.31$. When accounting for the changes in attention through β_3 and β_4 the implied attention parameters are $\hat{\gamma}_{\pi,L} = \frac{\hat{\beta}_{2,L} + \hat{\beta}_3 \cdot 2}{\hat{\beta}_{1,L} + \hat{\beta}_4 \cdot 2} = 0.21$ at $\pi_{t-1} = 2\%$ and $\hat{\gamma}_{\pi,H} = \frac{\hat{\beta}_{2,H} + \hat{\beta}_3 \cdot 5}{\hat{\beta}_{1,H} + \hat{\beta}_4 \cdot 5} = 0.34$ at $\pi_{t-1} = 5\%$. Overall, these estimated attention parameters are quantitatively close to the ones from my baseline estimation of 0.18 and 0.35. As furthermore the interaction terms are insignificant, these findings suggest that the threshold model of attention captures the data well.

What drives attention?   Expression (5) predicts that attention to inflation is higher when the cost of information is lower or when the perceived persistence or prior uncertainty are higher. Consistent with the first one—that the cost of information is lower—I show in Appendix A 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.[15] Korenok et al. (forthcoming) also consider news coverage of inflation and estimate a threshold between 3.77-3.94%, which is very close to the threshold I estimate. They also estimate attention directly, using Google search data, and estimate for the U.S. that once inflation exceeds the threshold of 3.55% that attention increases with inflation. One reason why I may find a slightly 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

---

[15]Consistent with this, Bracha and Tang (forthcoming) 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.

not available for the years before 2004.[16] One caveat of this channel, however, is that a decrease in the cost of information should reduce people's posterior uncertainty, which, however, Weber et al. (2023) show is not the case in the data.

Not only is news coverage of inflation higher in times of higher inflation, the perceived persistence, $\rho_{\pi,r}$, also tends to be higher when inflation is above the threshold. In my baseline estimation, for example, the estimated β_1 increases from 0.78 to 0.85 when inflation exceeds the threshold. Weber et al. (2023) also find that the perceived inflation persistence has increased with inflation. They also show, however, that in their RCTs treatment effects when providing information about future inflation should not change when the increase in attention is solely driven by changes in ρ_π. The empirical findings, however, show that they do change, thus, suggesting that the perceived persistence is not the key driver of attention.

The last possibility, that σ_π^2 is higher when inflation is above the threshold, is supported by the data. For example, I find that quarter-on-quarter inflation volatility in my sample increases from 2.04% (annualized) to 3.21%, and the volatility of inflation expectations from 0.78% to 2.42%. Consistent with that, Weber et al. (2023) also arrive at the conclusion that changes in σ_π^2 are a likely driver of changes in people's attention to inflation across inflation regimes.

3 Inflation Dynamics Across Attention Regimes

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 roughly doubled the inflationary effects of supply shocks in the subsequent inflation surge.

Supply shocks. As my empirical measure of supply shocks, I use the oil supply news shocks from Känzig (2021) for the period 1975M1-2023M12. 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.

---

[16]Fasani et al. (2024) estimate a threshold VAR to identify shocks to long-horizon inflation expectations (shocks to the Fed's inflation target). They also identify two inflation regimes and estimate the median threshold to be at an inflation rate of 3.9%, very similar to my findings.

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 3.91%. As an alternative indicator for the attention regimes, I also use Google Trends data, with the drawback that this data is only available since 2004. However, it will allow me to disentangle attention from other changes in the economy that tend to arise when inflation is higher (e.g., changes in the frequency of price adjustments, see, e.g., Blanco et al. (2024)). The Google 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 75th percentile.

Table 3: Shock properties across regimes

Regime	Mean	Standard deviation
Inflation as regime-defining variable		
High	-0.029	0.576
Low	0.015	0.567
Google Trends as regime-defining variable		
High	-0.011	0.786
Low	0.006	0.610

Notes: This table shows the mean and the standard deviation 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.

Table 3 lists the mean and the standard deviation of the shocks for the high-attention and low-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 3.91% 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. When using Google Trends as the regime-defining variable, the volatility of the supply shocks is slightly higher. In the model calibration, I will exploit this as one potential explanation for the increase in people's attention. That said, however, the following empirical results remain robust even when excluding shocks that are larger than one standard deviation. Thus, the following results are unlikely to be driven by differences in the shock series across regimes.[17]

The role of the attention regime 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):

$$y_{t+j} = \mathbb{1}_H \left(\alpha_j^H + \beta_j^H \varepsilon_t \right) + (1 - \mathbb{1}_H) \left(\alpha_j^L + \beta_j^L \varepsilon_t \right) + \Gamma' X_t + u_{t+j}, \tag{15}$$

---

[17]Excluding large shocks also helps to address the issues that may arise with state-dependent local projections when the state is endogenous (Gonçalves et al., 2024).

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 at the time of the shock 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, of inflation and of inflation expectations as controls. In Appendix C, I show that the results remain robust when using other controls.

Figure 2: Inflation response to an oil supply news shock

(a) High-Attention Regime

(b) Low-Attention Regime
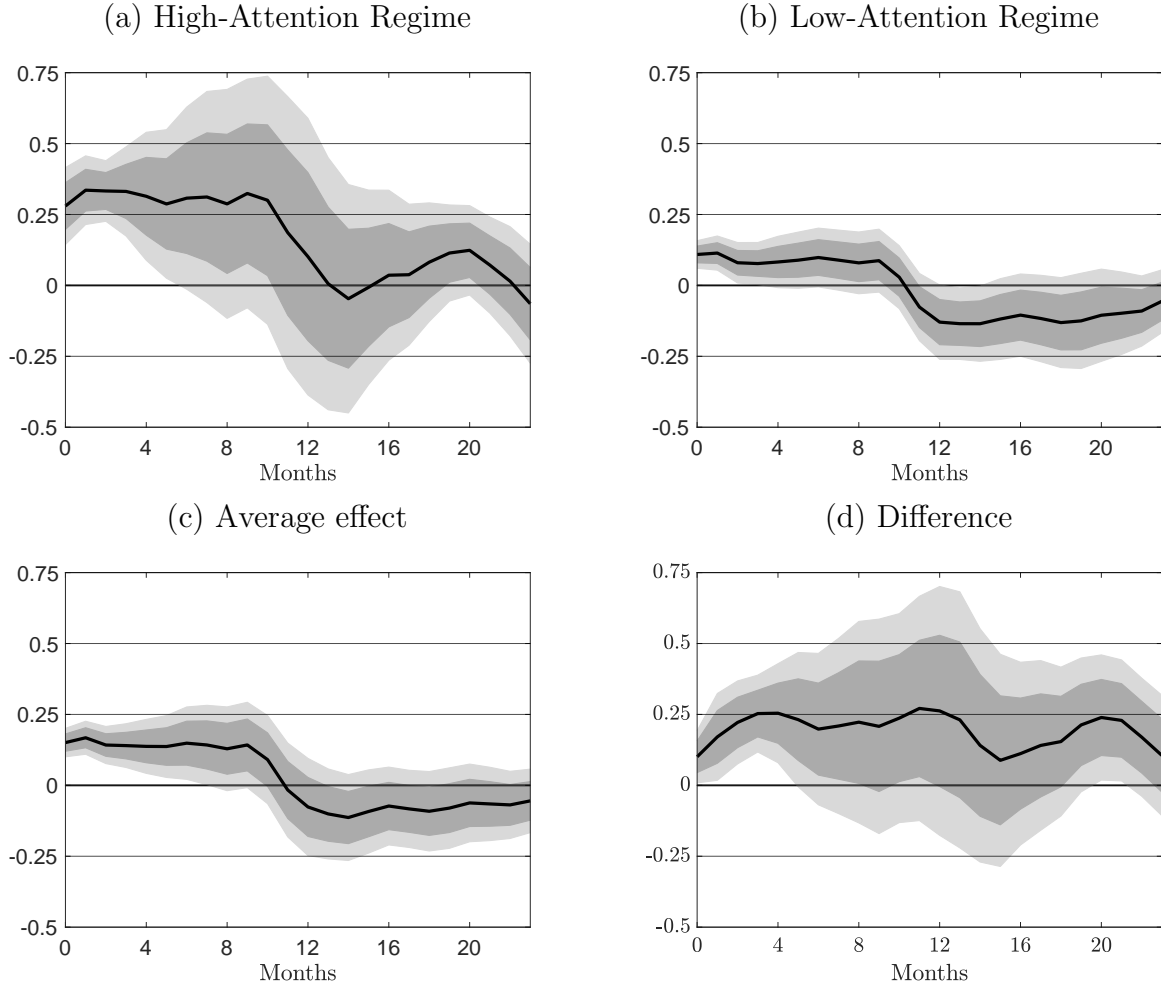
(c) Average effect
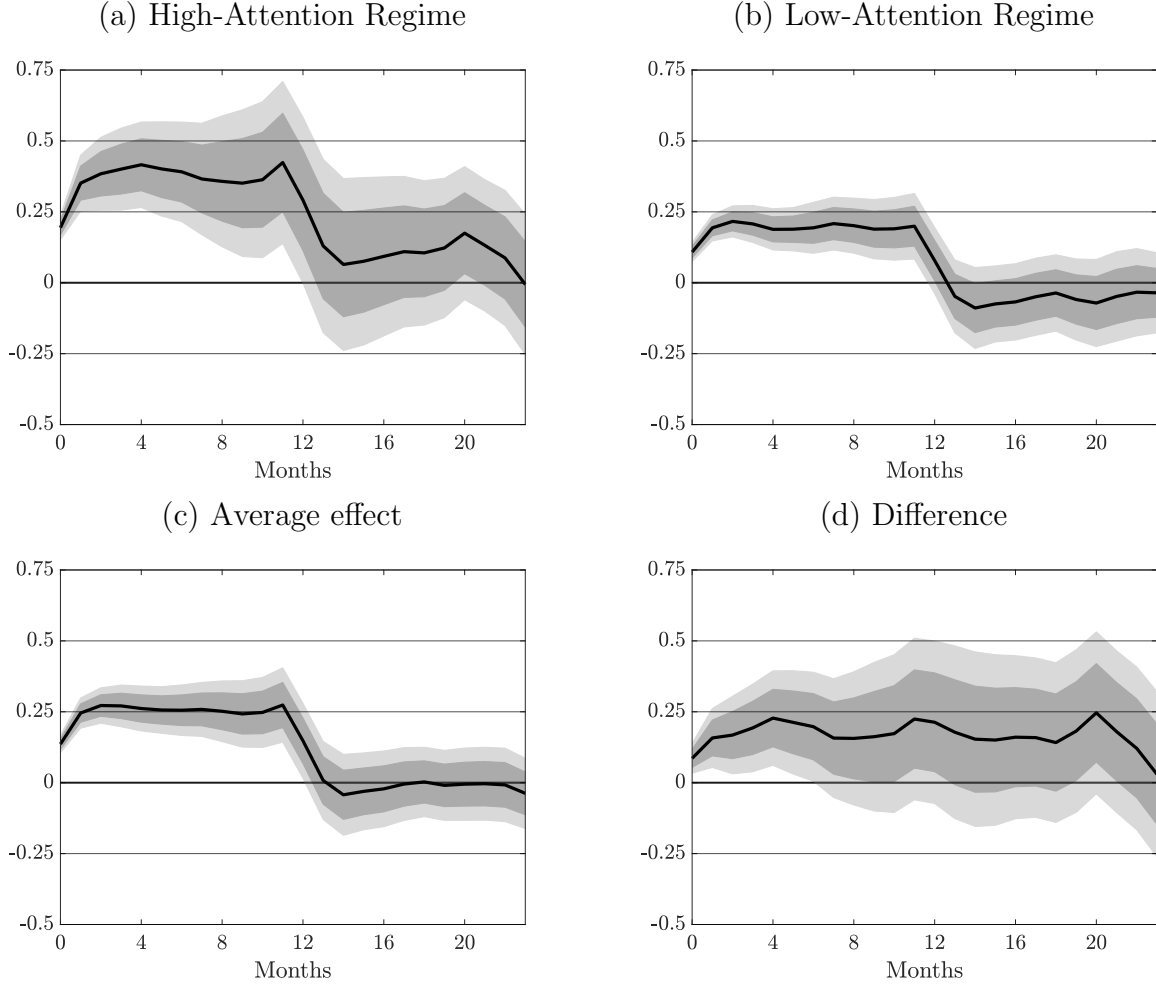
(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock (equation (15)) in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 2 shows the estimation results of regression (15). Panel (a) depicts the inflation response to 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).

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 40 basis points. In contrast, when attention to inflation is low at the time of the shock, inflation only increases by about 20 basis points. Figure 14 in Appendix C shows that the results are very similar when using Google Trends as the regime-defining variable.

Regional variation.   As in Section 2, I now also use regional variation of the four consensus regions. In particular, I use region-specific CPI inflation as the dependent variable. Additionally, I control for region fixed effects as well as four lags of region-specific inflation.

Figure 3: Inflation response to an oil supply shock using regional variation

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)) when using regional variation and region-specific CPI inflation. 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 heteroskedasticity as well as serial and cross-sectional correlation.

Figure 3 shows the results. We see that the results are very similar to my baseline specification

even though the overall effects are slightly smaller and the confidence bands for the high-attention regime wider.

The regional data also helps me to disentangle attention from other phenomena that arise with higher inflation. To do so, I take two approaches. In the first one, I use Google Trends as the attention-regime defining variable for the four different regions.[18] To control for other phenomena apart from attention that may arise with inflation, I further include an interaction term of the shock with lagged inflation, and I include region fixed effects. Panels (a)-(c) in Figure 4 show the results. In particular, we see that adverse oil supply news shocks have a strong and persistent effect on inflation when attention is high, whereas the effects in times of low attention are substantially weaker. In addition, the interaction term with inflation is practically zero at all horizons. These results therefore suggest that attention to inflation renders supply shocks more inflationary even when controlling for inflation itself. The results in panels (a) and (b) also suggest that the inflation response is *qualitatively* different across the two regimes. While inflation peaks on impact in the low-attention regime, it shows more of a hump-shaped response in the high-attention regime.

In the second approach, I control for time and region fixed effects and for regional-specific inflation. However, since the regimes based on Google Trends overlap strongly across regions—and are therefore largely absorbed by time fixed effects—I instead include an interaction term of last month's Google Trends searches with the shock rather than focusing on regimes. Panel (d) in Figure 4 shows that this interaction term is highly significant in the first 10 months (the coefficient is normalized to show the additional effect of a one standard deviation shock when Google Trends is one standard deviation above its mean). These results indicate that the different inflation responses across attention regimes are unlikely to be solely driven by aggregate phenomena that may confound the results, such as different policy responses across regimes, a general increase in the frequency of price adjustments, or an increase in the slope of the aggregate Phillips curve. Instead, the significant effects arising from more Google searches of inflation indicate that higher attention indeed drives at least partly the different inflation responses observed across regimes.

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 12, 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 13, 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). These findings suggest that higher attention to inflation not only renders supply shocks more inflationary, but shocks more generally become more inflationary. The predictions of the model derived later will be consistent with these findings.

---

[18]I take the average within each month across all states within a particular consensus region in order to aggregate Google Trends at the consensus region level.

Figure 4: Disentangling attention and inflation
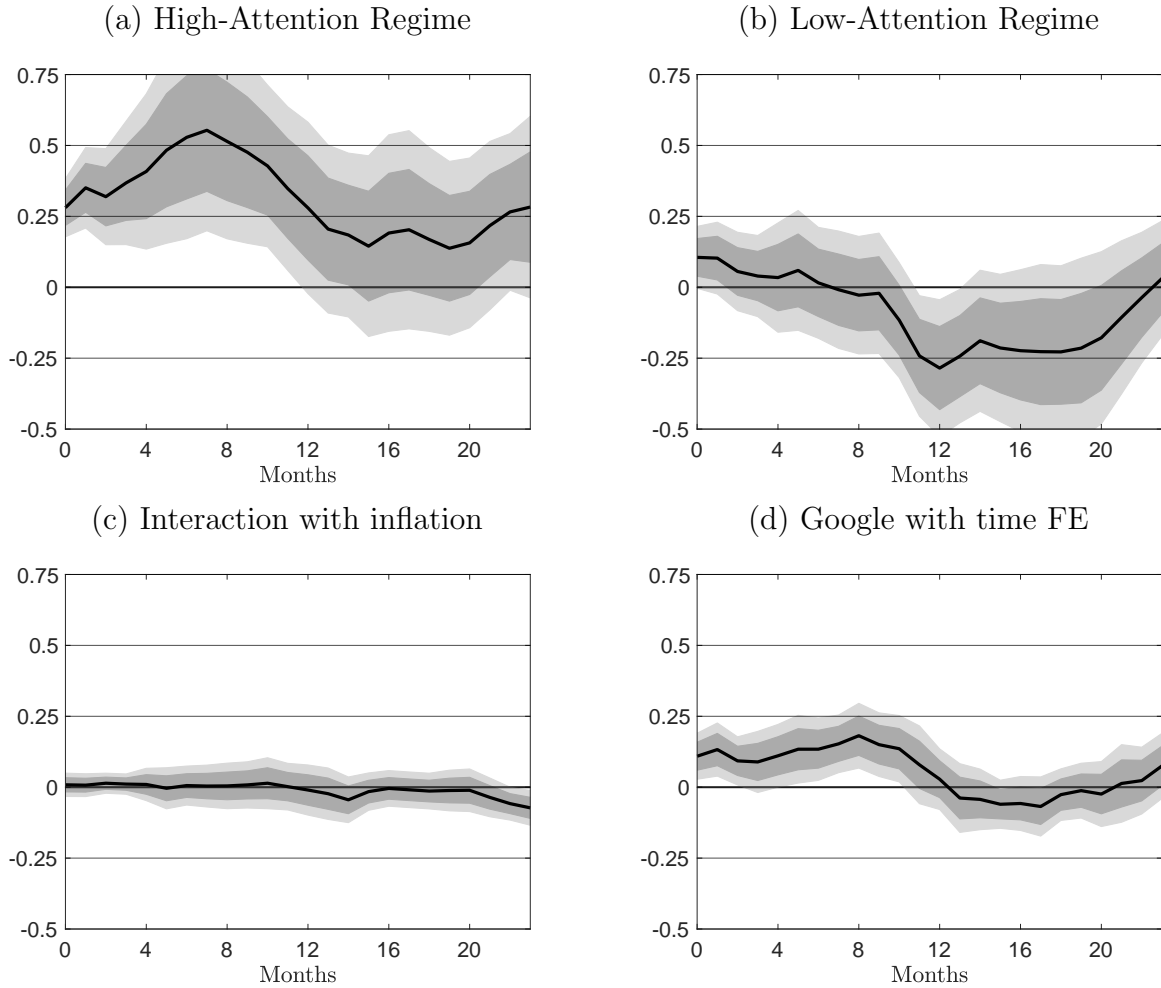
(a) High-Attention Regime

(b) Low-Attention Regime

(c) Interaction with inflation

(d) Google with time FE

Notes: This figure shows the inflation response to an oil supply news shock in the high-attention regime (panel (a)) and the low-attention regime (panel (b)) when using regional CPI data and when controlling for the interaction of the shock with last-month's inflation (which is shown in panel (c)). The attention regimes are based on Google Trends. Panel (d) shows the interaction term of Google Trends with the shock (for a one standard deviation shock and Google Trends being one standard deviation above its mean) when controlling for regional CPI inflation and time and region fixed effects. Standard errors are robust with respect to heteroskedasticity as well as serial and cross-sectional correlation.

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 15 and 16) or a higher sensitivity of macroeconomic variables in general (figures 17 and 18). The results are very similar when excluding the Covid period (figure 19), or when excluding the Great Inflation period (figure 20). Similarly, the results remain robust when using different control variables (figure 21), 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 22). 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 the increase in people's attention to inflation played a quantitatively important role in the post-Covid inflation surge in the United States.[19] 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 3.91% and thus, the economy entered the high-attention regime.[20] 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 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 almost 7%. Thus, the oil supply shocks account for about 55% 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.[21]

How much of this is due to the increase in attention? The red-dotted line shows that if there 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 1.5 percentage points (compared to the 2.7pp. 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.

To better understand the underlying mechanisms of these results as well as why the increase in people's attention may explain the interplay between inflation and inflation expectations (as shown in panel (b) in Figure 1), I now introduce the attention threshold and changing degrees of people's attention to inflation into an otherwise standard New Keynesian model.

---

[19]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.

[20]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 U.S. in mid-2021. Reis (2022) also documents that household inflation expectations in mid-2021 were not very well anchored any more.

[21]Even though I focus on supply shocks, the increase in people's attention to inflation in 2021 likely also increased the inflationary effects of demand shocks, such as monetary stimulus, given that I also find demand shocks to be more inflationary in times of higher attention (Figure 13).

Figure 5: Supply shocks, attention and the post-Covid inflation surge
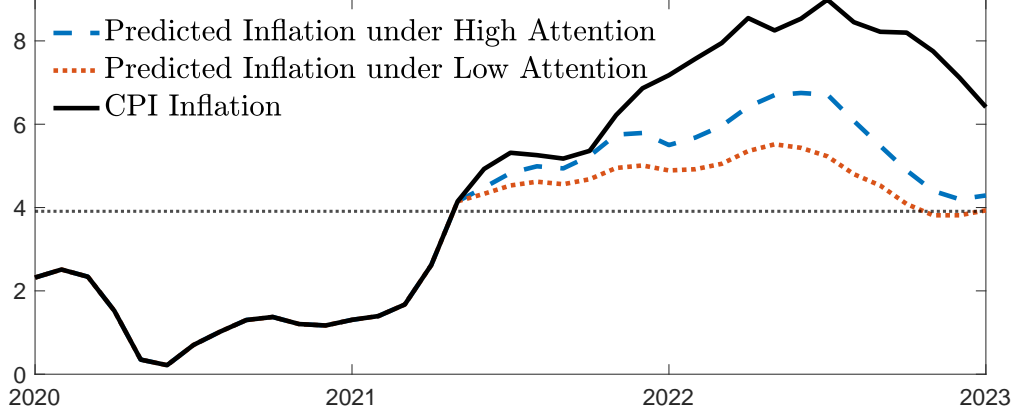
Notes: This figure shows the actual U.S. CPI inflation from 2020 until 2023, as well as the counterfactual inflation dynamics arising only from oil supply news shocks in the economy accounting for the increase in attention in April 2021 (blue-dashed line) and in the economy ignoring the increase in attention (orange-dotted line).

4 A Monetary Model with the Attention Threshold

In this section, I develop a monetary model which builds on the textbook New Keynesian model (Woodford, 2003; Galí, 2015) but households and firms pay only limited attention to inflation.

4.1 Households

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

$$\tilde{E}_0 \sum_{t=0}^{\infty} \beta^t Z_t \left[\log\left(C_t\right) - \Xi H_t \right], \tag{16}$$

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 parameter Ξ is the utility weight on hours worked.

Households maximize their lifetime utility subject to the flow budget constraints

$$C_t + B_t = w_t H_t + \frac{1 + i_{t-1}}{1 + \pi_t} B_{t-1} + \frac{T_t}{P_t}, \quad \text{for all } t, \tag{17}$$

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 (16) subject to (17) yields the Euler equation

$$Z_t C_t^{-1} = \beta(1 + i_t) \tilde{E}_t \left[Z_{t+1} C_{t+1}^{-1} \frac{1}{1 + \pi_{t+1}} \right], \tag{18}$$

and the labor-leisure condition

$$w_t = \Xi C_t. \tag{19}$$

The labor-leisure equation (19) shows that consumption fluctuations are solely driven by aggregate wage fluctuations.

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. (forthcoming).

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

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}, \tag{20}$$

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:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t. \tag{21}$$

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

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}. \tag{22}$$

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

$$Y_t(j) = H_t(j). \tag{23}$$

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.[22] These taxes are given back to firms in a lump-sum fashion, denoted $t_t^F(j)$. Taxes are assumed to be constant in the efficient economy, i.e., absent price rigidities, but fluctuate around their steady state in the economy with price rigidities in order to give rise to exogenous cost-push shocks.

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

---

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

(2011). I first discuss 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.[23] The per-period profits (in real terms) are given by

$$(1 - \tau_t)P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}\frac{Y_t}{P_t} - w_tH_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 Y_t + t_t^F(j), \tag{24}$$

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

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

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

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

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

$$\widehat{p}_t(j) = \frac{1}{\psi + \epsilon}\left[\psi\widehat{p}_{t-1}(j) + \epsilon\left(\widehat{mc}_t - \widehat{T}_t + \widehat{p}_t\right) + \beta\psi\tilde{E}_t^j\pi_{t+1}^j\right], \tag{25}$$

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 $\tilde{E}_t^j\pi_{t+1}^j$ or differences in past prices $\widehat{p}_{t-1}(j)$.

Government. The government imposes a sales tax τ_t on sales of intermediate goods, issues government 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

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

---

[23]Thus, I abstract from changes in the frequency of price adjustments. Cavallo et al. (forthcoming) and Blanco et al. (2024) argue that the frequency of price adjustments has increased during the recent inflation surge. While it is difficult to empirically disentangle attention to inflation from changes in price-setting behavior (which itself might change due to a change in attention of price setters), my results using Google Trends data as well as the survey evidence from Schwartzman and Waddell (2024) suggest that changes in price-setting behavior cannot explain the whole inflation dynamics and that attention indeed played an important role.

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

$$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1 - \rho_i)\left(\phi_\pi \pi_t + \phi_x \widehat{x}_t\right), \tag{26}$$

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

The subjective expectations of households and firm managers are modelled in such a way that there is a direct mapping from the empirical results of Section 2 to the model.

Households. Households have to form expectations about inflation and consumption.[24] According to the labor-leisure equation (19), consumption fluctuations are solely driven by aggregate wage fluctuations. Households therefore use information about aggregate wages to form their consumption expectations. Households believe that wages and inflation both follow a random walk:

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

where $\nu_{\pi,t}$ and $\nu_{w,t}$ are normally distributed with mean zero, and independent from each other. I discuss the case in which households hold rational expectations about consumption later. The random walk assumption ensures that long-run expectations align with the actual long-run variables.[25] They observe the aggregate shocks τ_t and Z_t perfectly.

At the time the household forms her expectations about future wages and inflation, she does not perfectly observe their current realizations, capturing that the household is not at all times perfectly monitoring all consumption expenditures and wage incomes 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

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

with normally distributed noise terms $\varepsilon_{\pi,r,t}$ and $\varepsilon_{w,t}$. I assume that signals are public signals and every agent in the economy receives the same signals. This implies that all agents have the same information and hence, hold the same expectations.

As detailed in Appendix D, potential output, i.e., output under flexible prices, is constant and

---

[24]I assume that bonds adjust residually such that the budget constraint always holds.
[25]Assuming an AR(1) with $\rho_\pi < 1$ does not qualitatively change the results.

consumption in log-deviations from steady state is equal to the output gap $\hat{x}_t$ in equilibrium (and equal to wages in deviations from steady state), $\hat{c}_t = \hat{w}_t = \hat{x}_t.$[26] It follows that if we assume initial values $\tilde{E}_{-1}\hat{c}_0 = \tilde{E}_{-1}\hat{w}_0 = \tilde{E}_{-1}\hat{x}_0$, we have $\tilde{E}_t\hat{c}_{t+1} = \tilde{E}_t\hat{w}_{t+1} = \tilde{E}_t\hat{x}_{t+1}$ for all t. This holds true, even if the household does not know that consumption and wages are equal to the output gap in equilibrium.

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

$$\tilde{E}_t\hat{x}_{t+1} = \tilde{E}_{t-1}\hat{x}_t + \gamma_x \left(\hat{x}_t - \tilde{E}_{t-1}\hat{x}_t \right), \tag{27}$$

where $\tilde{E}_t\hat{x}_{t+1}$ 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 $\tilde{E}_{t-1}\hat{x}_t$, whereas $\gamma_x = 1$ captures the case of full attention in which case the agent believes $\tilde{E}_t\hat{x}_{t+1} = \hat{x}_t$, 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 in the baseline case. I discuss the cases with regime-dependent γ_x and with rational consumption expectations later.

Inflation expectations follow the law of motion derived in equation (7) for $\rho_\pi = 1$, so that they are given by

$$\tilde{E}_t\pi_{t+1} = \begin{cases} \tilde{E}_{t-1}\pi_t + \gamma_{\pi,L} \left(\pi_t - \tilde{E}_{t-1}\pi_t \right), & \text{when } \pi_{t-1} < \bar{\pi} \\ \tilde{E}_{t-1}\pi_t + \gamma_{\pi,H} \left(\pi_t - \tilde{E}_{t-1}\pi_t \right), & \text{when } \pi_{t-1} \geq \bar{\pi}, \end{cases} \tag{28}$$

where $\gamma_{\pi,r}$ captures the optimal level of attention to inflation in regime $r \in \{L, H\}$.[27] I leverage the expression (5) for optimal attention to inflation

$$\gamma_{\pi,r} = max \left\{ 0, 1 - \frac{\frac{1}{\lambda_r}}{2\chi\rho_{\pi,r}^2\sigma_{\pi,r}^2} \right\}$$

to calibrate the model. Defining the cost of information relative to the stakes χ, $\tilde{\lambda} \equiv \frac{1}{2\lambda\chi}$ and imposing $\rho_{\pi,r} = 1$, I obtain

$$\gamma_{\pi,r} = max \left\{ 0, 1 - \frac{\tilde{\lambda}_r}{\sigma_{\pi,r}^2} \right\}. \tag{29}$$

As my proxy for $\sigma_{\pi,L}^2$, I use the model-implied variance of inflation in regime L given the rest of the calibration (discussed in Section 4.5). I set $\tilde{\lambda}$ such that $\gamma_{\pi,L} = 0.18$ as found in the data. In order to pin down $\gamma_{\pi,H}$, I then increase the exogenous volatility of the cost-push shocks whenever inflation

---

[26]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 $\hat{y}_t = \hat{c}_t$.

[27]Without loss of generality, I do not explicitly model noise shocks. Accounting for noise shocks would introduce an additional exogenous shock, but does not affect my analysis of how supply and demand shocks are transmitted.

exceeded the attention threshold in the previous period. I then solve for the fixed point such that $\gamma_{\pi,H} = 0.36$ and that the model-implied variance of inflation is equal to the $\sigma^2_{\pi,H}$ implied by equation (29) for $\gamma_{\pi,H} = 0.36$ and the cost of information $\tilde{\lambda}$ obtained before.

As an alternative, the attention change may arise due to a change in the exogenous cost of information, $\tilde{\lambda}_H < \tilde{\lambda}_L$, which could reflect the fact that news coverage of inflation is substantially higher in times of high inflation, thus, lowering the cost of information in these periods. All the following results are very similar under that alternative assumption.

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., $\tilde{E}_t^j \pi_{t+1}(j) = \tilde{E}_t^j \pi_{t+1}$, and that they form expectations about inflation in the same way as households. Consistent with these assumptions, Yotzov et al. (2024) show that firms' expected own-price growth is strongly positively correlated with changes in CPI inflation, and McClure et al. (forthcoming) who show that managers and non-managers hold similar average inflation and unemployment expectations and respond similarly to information treatments.[28] As for households, all forecasters receive the same public signal about current inflation, from which it follows that $\tilde{E}_t^j \pi_{t+1}(j) = \tilde{E}_t^j \pi_{t+1} = \tilde{E}_t \pi_{t+1}$.

Assuming that firms all start out with the same initial price, i.e., $\hat{p}_{-1}(j) = \hat{p}_{-1}$ for all $j \in [0,1]$, it follows 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 (25) can be written as

$$\pi_t = \frac{\epsilon}{\psi}\left(\widehat{mc}_t - \widehat{T}_t\right) + \beta \tilde{E}_t \pi_{t+1}. \tag{30}$$

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

$$\pi_t = \beta \tilde{E}_t \pi_{t+1} + \kappa \hat{x}_t + u_t. \tag{31}$$

---

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

4.4 Equilibrium

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

$$\pi_t = \beta \tilde{E}_t \pi_{t+1} + \kappa \widehat{x}_t + u_t, \tag{32}$$

$$\widehat{x}_t = \tilde{E}_t \widehat{x}_{t+1} - \left(\tilde{i}_t - \tilde{E}_t \pi_{t+1} - r_t^* \right), \tag{33}$$

$$\tilde{i}_t = \rho_i \tilde{i}_{t-1} + (1 - \rho_i) \left(\phi_\pi \pi_t + \phi_x \widehat{x}_t \right), \tag{34}$$

as well as the two law of motions for output gap expectations and inflation expectations, equations (27) and (28). Equations (32)-(34) are identical to the 3-equation representation of the standard New Keynesian model (see Galí (2015)) except for the expectations formation of agents.

Equation (32) is the New Keynesian Phillips curve, representing the supply side of the economy, and equation (33) denotes the aggregate Euler (or IS) equation, which together with monetary policy (equation (34)) pins down aggregate demand. The natural interest rate r_t^* 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.

4.5 Calibration

Most of the calibration is standard. I set the discount factor β to target a steady state natural rate of 1% (annualized), 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 $\phi_x = 0.125$.

I assume that both shocks follow an AR(1) process with persistence 0.8. 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 the exogenous shock volatility when inflation is below the threshold, $\sigma_{u,L}$. This results in $\sigma_{u,L} = 0.3\%$. I assume that demand shocks have the same volatility, $\sigma_r = \sigma_{u,L}$ and I assume that the demand shock volatility is the same in both regimes. Given my empirical estimates, I set $\gamma_{\pi,L} = 0.18$ and the threshold to $\bar{\pi} = 3.91\%$ (annualized). In order to calibrate the cost-push shock volatility in the high-attention regime $\sigma_{u,H}$, I solve for the fixed-point problem discussed in Section 4.3 by targeting $\gamma_{\pi,H} = 0.35$. To do so, I simulate the model for 10,000 periods and vary $\sigma_{u,H}$ until the fixed point is reached. This results in $\sigma_{u,H} = 1.23 \cdot \sigma_{u,L}$. That said, however, the subsequent results are very similar when imposing that the change in attention arises from a change in the cost of information and therefore, keeping the shock volatility constant across regimes.

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 D.1). This

results in $\gamma_x = 0.25$ for both regimes, and hence, I impose that γ_x does not change across regimes and set it to $\gamma_x = 0.25$.

As a sanity check of the proposed empirical approach to estimate the attention threshold and the degrees of attention in a setup in which inflation and inflation expectations are determined jointly and different type of shocks are hitting the economy, I estimate the threshold regression (8) on model-simulated data and find that the regression results are exactly equal to the calibrated parameters.

5   Inflation Surges

In this section, I show how the model with two inflation-attention regimes can jointly generate persistent heightened inflation periods, a long last mile of disinflation, forecast error dynamics mirroring the ones recently observed in the United States, and lead to an asymmetry in the dynamics of inflation. In Appendix E.1, I show in a slightly stylized version of the model how the model can be represented in a standard AS-AD framework in which crossing the attention threshold leads to a steepening of both the AS and the AD curve. In 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. Additionally, the heightened attention amplifies the inflationary effects of shocks.
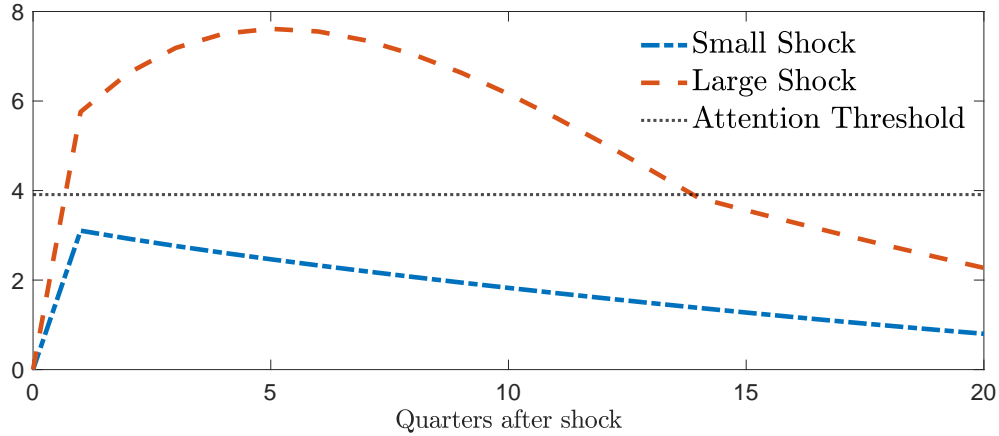
5.1   Crossing the attention threshold

Equipped with these theoretical insights, I now move to a numerical analysis of the attention threshold in which I relax the simplifying assumptions made in the previous section and relying on the calibration discussed in section 4.5. Figure 6 shows the inflation dynamics following a cost-push shock that pushes inflation in the first period above the attention threshold $\bar{\pi}$ with a persistence of 0.8 (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).

The two cases exhibit fundamentally different inflation dynamics. When attention remains low, which happens in response to the relatively small shock, inflation peaks 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. These self-reinforcing dynamics in the first few periods would not be present in the case of deflationary shocks because attention to inflation stays constant in that scenario. Therefore, the model with the attention threshold offers a simultaneous explanation for the recent inflation surge as well as the 'missing deflation puzzle' in the aftermath of the Great Financial crisis (see, e.g., Coibion and Gorodnichenko (2015b)). Additionally, the different inflation dynamics also mirror the ones discussed in Figure 4, namely that inflation peaks on impact

when attention is low whereas it is hump shaped in times of high attention.[29]

Figure 6: Inflation dynamics

Notes: The blue-dashed-dotted line shows the (annualized) inflation response to a cost-push shock that does not push inflation above the attention threshold, and the red-dashed line shows the inflation dynamics for a cost-push shock that is large enough to push inflation above the threshold.

To understand the inflation surge 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 are 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. 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. This initial period of disinflation happens relatively quickly. The reason is that once inflation actually starts to come down, households and firm managers are still highly attentive and hence, revise their expectations downwards strongly. Thus, actual inflation falls relatively quickly.

However, once inflation falls back below the threshold $\bar{\pi}$—and the second regime switch occurs—the speed of disinflation decreases. 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. Their priors are now relatively high because of the high inflation period that agents just lived through, and therefore, their inflation expectations remain persistently high which hinders actual inflation from decreasing quickly.[30]

---

[29]Figure 26 shows the response of inflation when the shock hits in times of high-attention and illustrates that the inflation response is indeed hump shaped in that case.

[30]Lebow and Peneva (2024) also show that the perceptions of inflation among households came down more slowly than actual inflation.

Figure 7: Inflation and inflation expectation dynamics

Model

Data
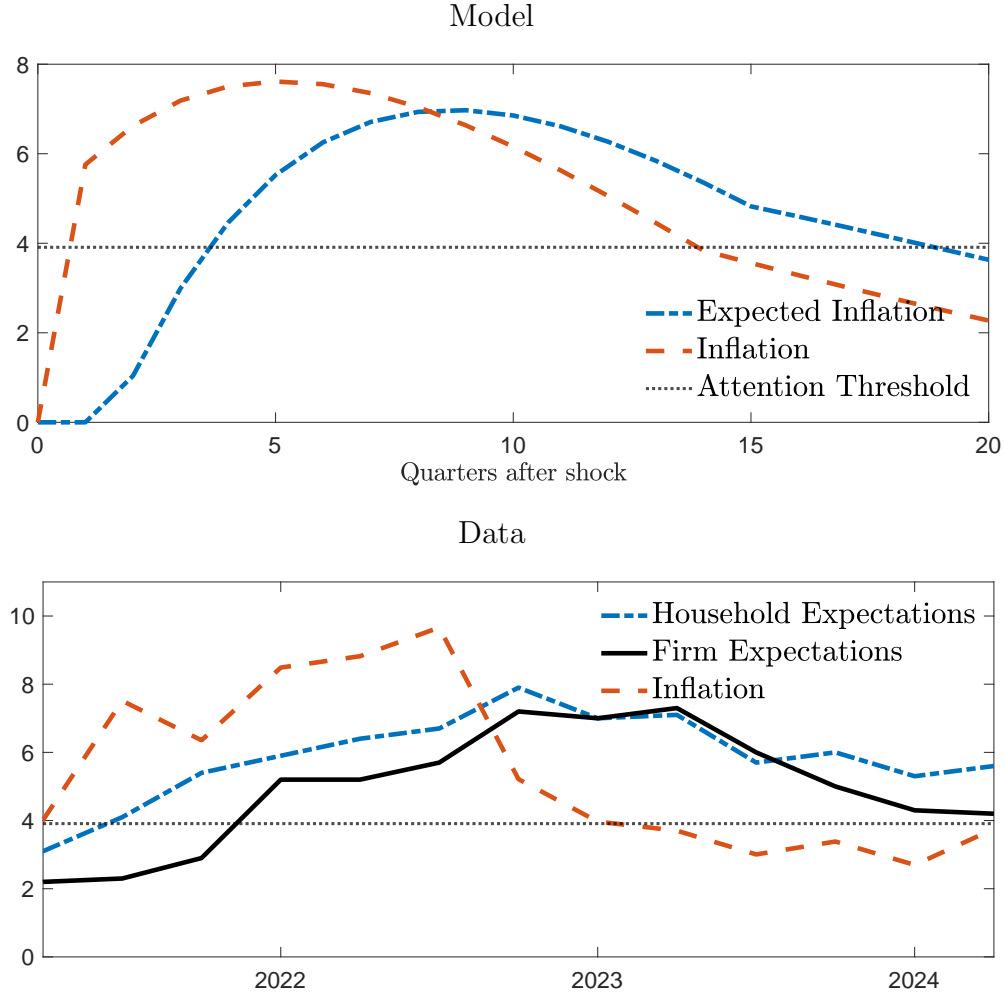
Notes: The upper panel shows the dynamics of inflation and inflation expectations after a cost-push shock that pushes inflation above the attention threshold. The lower panel plots quarter-on-quarter (annualized) CPI inflation and average inflation expectations for the period 2021Q1-2024Q1, as well as the estimated attention threshold at 3.91%. In both panels, the inflation expectations are shifted such that the vertical distance between the two shows the respective forecast errors.

To clearly see these patterns of inflation and inflation expectations, the upper panel in figure 7 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 inflation surpasses actual inflation. Hence, forecast errors become negative.

These patterns are consistent with the behavior of inflation and inflation expectations during the recent inflationary period, as documented in panel (b) in Figure 1 in the Introduction. The lower panel in figure 7 shows the same time series but for quarter-on-quarter inflation to make them consistent with the model (Figure 23 in Appendix C shows these patterns in the data conditional

on an oil supply news shock).[31] The figure shows annualized quarter-on-quarter CPI inflation (red-dashed line) and average inflation expectations from the Michigan Survey (blue-dashed-dotted line) and from the Survey of Firms' Inflation Expectations (black-solid line) for the period from 2021Q1 until 2024Q1. Consistent with the model, inflation peaks about a year and a half after exceeding the threshold. 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 two years after inflation exceeded the attention threshold. Consistent with 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 may lead to a prolonged period of inflation stubbornly above the central bank's target even though the initial dis-inflationary phase happens relatively quickly. Put differently, once attention to inflation decreases again, the higher prior expectations after the inflation surge may render the 'last mile' of inflation back to target more arduous. This also implies that the persistently higher inflation expectations give 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.

In Appendix E.2, I show that these inflation dynamics are unique to the model with the attention threshold. In particular, the models under full-information rational expectations (FIRE) or in which attention remains low predict that inflation peaks on impact and then comes down relatively quickly. Additionally, models with FIRE could not account for the observed dynamics of forecast errors. A model variant in which attention to inflation is always high, on the other hand, can produce the initial hump-shaped response but then predicts a fast decline back to target, i.e., this model would not explain why we see such a long last mile of disinflation. Furthermore, a model in which attention is always high would have predicted a very pronounced deflationary period after the Great Financial crisis.

In Appendix E.2, I further discuss (i) the inflation dynamics following a demand shock (Figure 27), (ii) what the model predicts when attention to the output gap decreases in times of high attention to inflation (Figure 28), and (iii) how the inflation dynamics look like when consumption expectations are formed under FIRE. In cases (i) and (ii), the inflation surge is somewhat more persistent but otherwise very similar to the baseline case of the discussed cost-push shock with constant attention to the output gap, whereas in (iii) the model produces slightly less endogenous persistence.

---

[31]Angeletos et al. (2020b) show that this initial undershooting followed by a delayed overshooting of expectations is a general pattern in response to shocks in the data.

5.2 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 see long-lived periods of pronounced deflation after such a shock, consistent with the fact the U.S. economy did not experience any long lasting deflationary periods since the second World War. These asymmetries matter for the overall properties of inflation. In particular, ignoring the attention threshold results in underpredicting the risk of persistently high inflation rates.

Figure 8: The attention threshold and inflation asymmetry
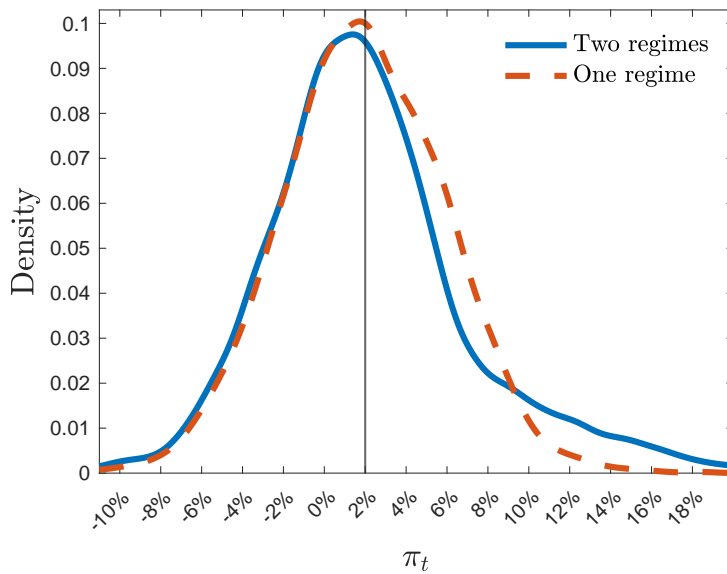
Notes: This figure shows the distribution of inflation (annualized) for the model with the attention threshold (solid-blue line) and the one without the attention threshold (dashed-red line).

Figure 8 plots the distribution of inflation for the model with the attention threshold (blue-solid line) and for the limited-attention model without the threshold (red-dashed line). These figures are obtained by simulating the model for 10,000 periods by hitting it with random normally-distributed cost-push and demand shocks.

The model with the attention threshold has a substantially thicker right tail than the one without the attention threshold, as is the case empirically.[32] The probability that inflation exceeds 8% (annualized) is about 8.2% in the model with the threshold, and thus substantially closer to its empirical counterpart of 12% than in the model without the threshold in which inflation exceeds 8% only 2% of the time. When it comes to periods of deflation, the two models exhibit very similar properties. For example, both models predict inflation to fall below -8% in only 2-3% of the periods,

---

[32]Note, that even though the volatility of cost-push shocks is slightly higher in the higher-attention regime, the asymmetry also arises when shocks are drawn from the same distribution in both regimes due to the attention threshold.

compared to 0.5% empirically, and the median in both models is very close to 0.[33] 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.

6 Monetary Policy 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 the central bank's loss function as

$$-\frac{1}{2}E_0 \sum_{t=0}^{\infty} \beta^t \left[\pi_t^2 + \Lambda \widehat{x}_t^2\right], \tag{35}$$

where Λ is the relative weight of the output gap, which I set to $\Lambda = 0.007$ as in Adam and Billi (2006). The policymaker wants to maximize (35) subject to the private sector's optimality conditions, characterized by equations (32) and (33). But I assume that the planner wrongly believes that the private sector holds rational expectations. The problem from the policymaker's perspective is then identical to the one in Clarida et al. (1999) who show that the optimal policy is given by

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

if the policymaker can fully commit to its policy. If the policymaker cannot commit, the optimal policy rule under discretion is given by

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

In the following, I compare the implications of these two policy rules to the Taylor rule (34) as well as to a Taylor rule without interest rate smoothing and no response to the output gap, i.e., to (34) 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	$\widetilde{i}_t = \rho_i \widetilde{i}_{t-1} + (1 - \rho_i)\left(\phi_\pi \pi_t + \phi_x \widehat{x}_t\right)$
(2)	Taylor rule without smoothing	$\widetilde{i}_t = \phi_\pi \pi_t$
(3)	Optimal RE commitment policy	$\pi_t + \frac{\Lambda}{\kappa}\left(\widehat{x}_t - \widehat{x}_{t-1}\right) = 0$
(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. Panel (a)

---

[33]The mean is 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.

in Figure 9 plots the central bank's losses (35) 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 (b) shows the respective inflation volatilities, panel (c) the average level of inflation, and panel (d) the frequency of how often inflation is above the threshold of 4%.

Figure 9: Implications of different monetary policy rules

(a) Central bank loss

(b) Inflation volatility   (c) Average inflation   (d) Frequency high-attention
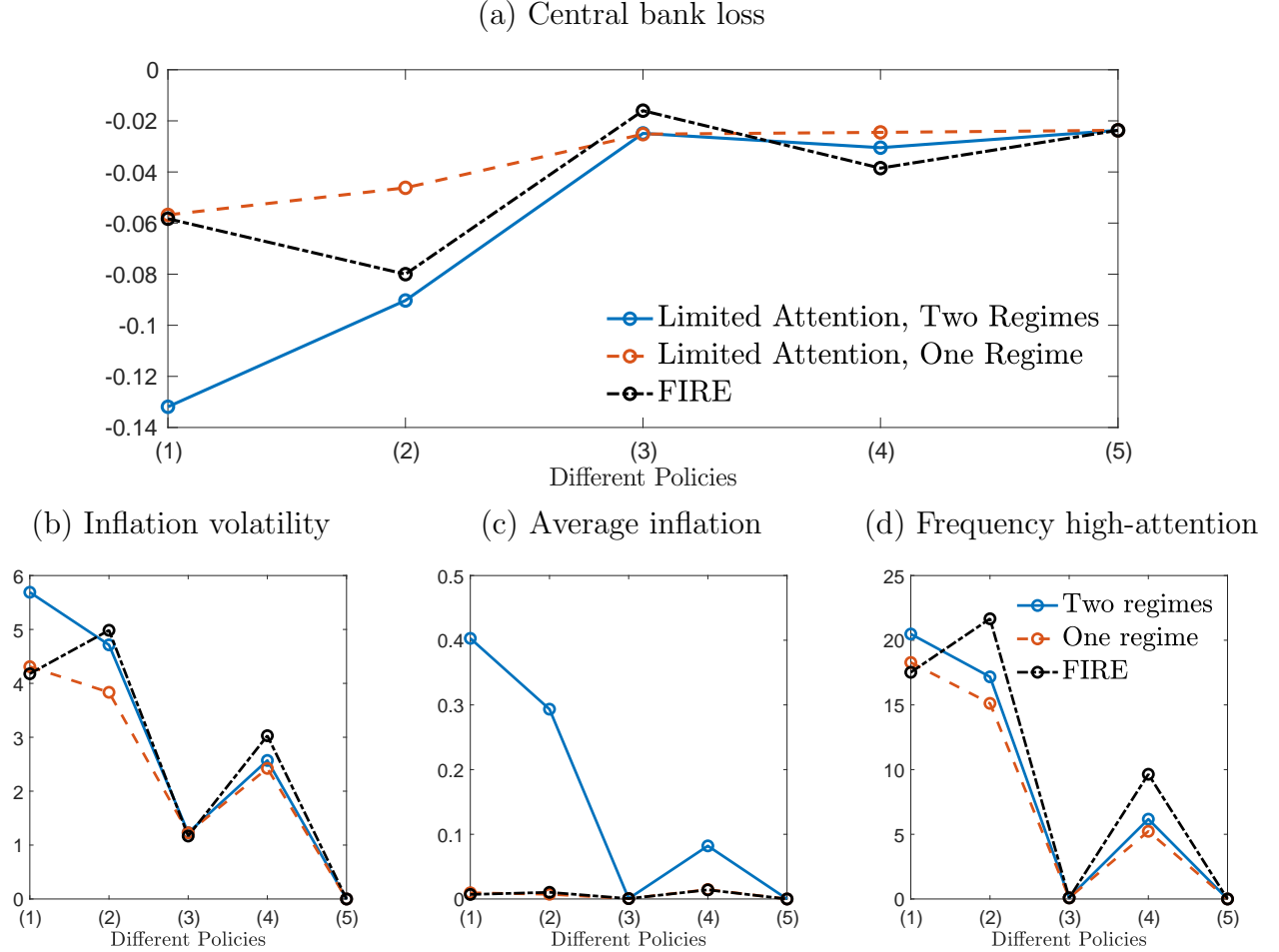
Notes: This figure shows the central bank loss (using equation (35) and shown in panel (a)), inflation volatility (panel (b)), average inflation (panel (c)), and the frequency of being in the high-attention regime (panel (d)) for the five different monetary policy rules from Table 4.

There are two main takeaways from Figure 9. First, following simple Taylor rules (policy rules (1) and (2)) leads to substantially larger central bank 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.[34] 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 (d) in figure 9). Due to the asymmetry of the attention threshold the average level of inflation is higher when the economy

---

[34]Gáti (2023) shows that in her model with varying degrees of long-run-expectations anchoring, 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. Similarly, Benchimol and Bounader (2023) find that under inattention to prices, central bank losses under Taylor rules are higher than with other policy rules.

is in the high-attention regime frequently (panel (c)). This is 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 these losses can be mitigated when monetary policy follows one of the rules (3)-(5). In these cases, inflation is relatively stable and fluctuates almost symmetrically around 0, as the economy very rarely stays in the high-attention regime. 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 losses are very similar to the ones under the other policy rules from (3) and (4).

Overall, Figure 9 illustrates that following simple Taylor rules, especially ones with interest-rate smoothing and relatively low inflation-response coefficients, can lead to large central bank losses when the public's attention to inflation increases 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 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 about 4% 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 effects 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 U.S. As inflation exceeds the threshold, the increase in people's attention leads to an endogenous amplification of the initial inflation surge. After inflation peaks, inflation comes back down relatively quickly initially, but once attention to inflation decreases again, the higher prior expectations after the inflation surge may render the 'last mile' of inflation back to target more arduous.

Accounting for varying attention levels also matters for the model's normative predictions. I show that that following simple Taylor rules leads to substantially larger central bank losses compared to more hawkish rules. While I do not characterize fully optimal monetary policy when people's attention to inflation is subject to sudden changes, doing so and studying how the conduct of monetary policy feeds back into people's attention is a promising direction for future research.

References

AASTVEIT, K. A., H. C. BJØRNLAND, AND J. L. CROSS (2023): "Inflation expectations and the pass-through of oil prices," *Review of Economics and Statistics*, 105, 733–743.

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 (2021): "Dynamic inattention, the Phillips curve, and forward guidance," *Working Paper*.

AMITI, M., S. HEISE, F. KARAHAN, AND A. ŞAHIN (2023): "Inflation Strikes Back: The Role of Import Competition and the Labor Market," *NBER Macroeconomics Annual, Volume 38*.

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 (forthcoming): "Understanding US inflation during the covid era," *Brookings Papers on Economic Activity*.

BAYER, C., B. BORN, AND R. LUETTICKE (forthcoming): "Shocks, frictions, and inequality in US business cycles," *American Economic Review*.

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

BENCHIMOL, J. AND L. BOUNADER (2023): "Optimal monetary policy under bounded rationality," *Journal of Financial Stability*, 67, 101151.

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*.

BERNANKE, B. AND O. BLANCHARD (2023): "What caused the US pandemic-era inflation?" *Hutchins Center Working Papers*.

BHANDARI, A., J. BOROVIČKA, AND P. HO (forthcoming): "Survey data and subjective beliefs in business cycle models," *Review of Economic Studies*,.

BIANCHI, F., R. FACCINI, AND L. MELOSI (2023): "A Fiscal Theory of Persistent Inflation," *The Quarterly Journal of Economics*.

BIANCHI, F. AND L. MELOSI (2022): "Inflation as a fiscal limit," *2022 Jackson Hole Symposium Volume*.

BINDER, C. AND R. KAMDAR (2022): "Expected and realized inflation in historical perspective," *Journal of Economic Perspectives*, 36, 131–155.

BLANCO, A., C. BOAR, C. J. JONES, AND V. MIDRIGAN (2024): "The Inflation Accelerator," *Working paper*.

BLANCO, A., P. OTTONELLO, AND T. RANOSOVA (2022): "The Dynamics of Large Inflation Surges," *Working paper*.

BRACHA, A. AND J. TANG (forthcoming): "Inflation levels and (in) attention," *The Review of Economic Studies*.

BROER, T., A. KOHLHAS, K. MITMAN, AND K. SCHLAFMANN (2023): "Information and wealth heterogeneity in the macroeconomy," *Working paper*.

BUELENS, C. (2024): "Googling "Inflation": What does Internet Search Behaviour Reveal about Household (In)attention to Inflation and Monetary Policy?" *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 (forthcoming): "The Inflation Expectations of US Firms: Evidence from a New Survey," *Journal of Monetary Economics*.

CANDIA, B., M. WEBER, Y. GORODNICHENKO, AND O. COIBION (2023): "Perceived and Expected Rates of Inflation of US Firms," *AEA Papers and Proceedings*, 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.

CAVALLO, A., F. LIPPI, AND K. MIYAHARA (forthcoming): "Large shocks travel fast," *American Economic Review: Insights*.

CERRATO, A. AND G. GITTI (2022): "Inflation since covid: Demand or supply," *Working paper*.

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

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

COIBION, O., Y. GORODNICHENKO, E. S. I. KNOTEK, AND R. SCHOENLE (2023): "Average Inflation Targeting and Household Expectations," *Journal of Political Economy: Macroeconomics*.

CRUMP, R. K., S. EUSEPI, M. GIANNONI, AND A. ŞAHIN (2024): "The unemployment-inflation trade-off revisited: The Phillips Curve in COVID times," *Working Paper*.

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.

ERCEG, C., J. LINDÉ, AND M. TRABANDT (2024): "Monetary Policy and Inflation Scares," .

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

FASANI, S., M. S. MIESCU, AND L. ROSSI (2024): "Nonlinearities With De-anchored Inflation Expectations," *Working Paper*.

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. (2023): "Monetary Policy & Anchored Expectations An Endogenous Gain Learning Model," *Journal of Monetary Economics*.

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

GONÇALVES, S., A. M. HERRERA, L. KILIAN, AND E. PESAVENTO (2024): "State-dependent local projections," *Journal of Econometrics*, 105702.

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.

GUILLOCHON, J. AND S. TER ELLEN (2023): "Inflation (In)Attention, Media, and Central Bank Trust," *Working Paper*.

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

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?" *Working Paper*.

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

JO, Y. J. AND S. ZUBAIRY (forthcoming): "State dependent government spending multipliers: Downward nominal wage rigidity and sources of business cycle fluctuations," *American Economic Journal: Macroeconomics*.

JONUNG, L. (1981): "Perceived and expected rates of inflation in Sweden," *The American Economic Review*, 71, 961–968.

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," *Working paper*.

KAMDAR, R. AND W. RAY (2024): "Attention-Driven Sentiment and the Business Cycle," *Working Paper*.

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 (forthcoming): "Inflation and attention thresholds," *Review of Economics and Statistics*.

KRONER, N. (2024): "Inflation and Attention: Evidence from the Market Reaction to Macro Announcements," *Working paper*.

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

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

Lebow, D. and E. Peneva (2024): "Inflation Perceptions During the Covid Pandemic and Recovery," *Working paper*.

Lorenzoni, G. (2009): "A theory of demand shocks," *American economic review*, 99, 2050–2084.

Lorenzoni, G. and I. Werning (2023): "Wage price spirals," *Forthcoming, Brookings Papers in Economic Activity*.

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.

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 (forthcoming): "The Macroeconomic Expectations of US Managers," *Journal of Money, Credit and Banking*.

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

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," *Working Paper*.

Pfäuti, O. (forthcoming): "Inflation–who cares? Monetary Policy in Times of Low Attention," *Journal of Money, Credit and Banking*.

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): *The burst of high inflation in 2021-22: How and why did we get here?*, Bank for International Settlements, Monetary and Economic Department.

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*.

SCHWARTZMAN, F. AND S. R. WADDELL (2024): "Inflation Expectations and Price Setting Among Fifth District Firms," *Richmond Fed Economic Brief*, 24.

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.

TUREN, J. (2023): "State-dependent attention and pricing decisions," *American Economic Journal: Macroeconomics*, 15, 161–189.

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

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," *Working Paper*.

WERNING, I. (2022): "Expectations and the Rate of Inflation," *Working Paper*.

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

YORK, J. (2023): "Do Household Inflation Expectations Respond to Macroeconomic Data Releases?" *Working Paper*.

YOTZOV, I., N. BLOOM, P. BUNN, P. MIZEN, AND G. THWAITES (2024): "The Speed of Firm Response to Inflation," *Working Paper*.

Online Appendix

A News Coverage of Inflation

Figure 10 shows 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).

Figure 10: News coverage of inflation is higher in times of high inflation
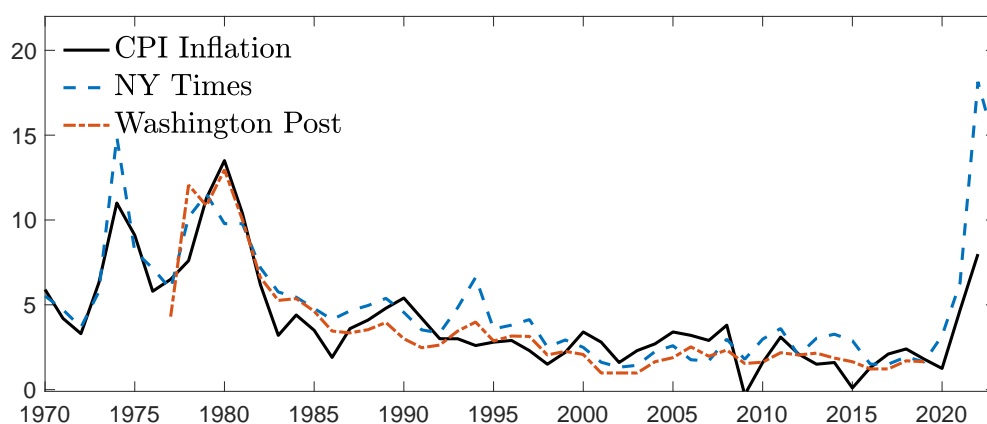
Notes: The black-solid line shows the annual CPI inflation rate for the U.S. since 1970, the blue-dashed line and the red-dashed-dotted line show the frequency of the word inflation in the New York Times and the Washington post, respectively.

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

Consistent with this, Bracha and Tang (forthcoming) 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. Guillochon and ter Ellen (2023) find that more exposure to inflation news leads to an increase in the probability that households pay attention to inflation. 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. Figure 11 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 a 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 seems to lag inflation slightly (the correlation between news coverage and *lagged* inflation is also 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 11: News coverage of inflation during Covid
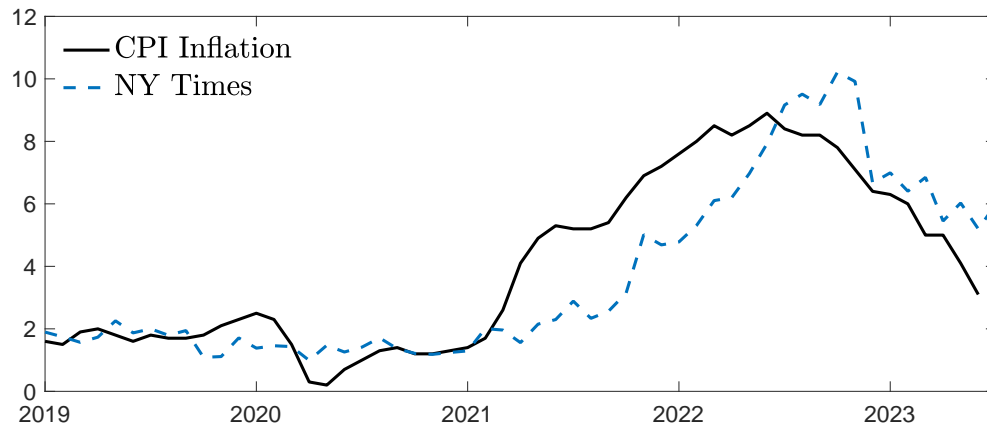
Notes: The black-solid line shows the monthly year-on-year CPI inflation rate for the U.S. and the blue-dashed line the frequency of the word inflation in the New York Times, normalized to have the same standard deviation as CPI inflation.

B   Appendix to Empirical Section 2

One-year-ahead expectations.   As my baseline in Section 2, I focus one one-quarter-ahead expectations and compare them to quarter-on-quarter inflation.   When using one-year-ahead expectations and comparing them to year-on-year inflation instead (and using last month's y-o-y inflation), I estimate an attention threshold at 5.03% and that attention increases from 0.54 to 1.22 from the low-attention to the high-attention regime.   The difference between the two is highly statistically significant with an associated *p*-value of 0.000.

Heterogeneity.   In the main text, I 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.[35]   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).

---

[35]There is a vast literature documenting heterogeneity in inflation expectations, see, e.g., D'Acunto et al. (2019); Broer et al. (2023); 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.

# C Additional Results and Robustness 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 their 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 12 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 significant.
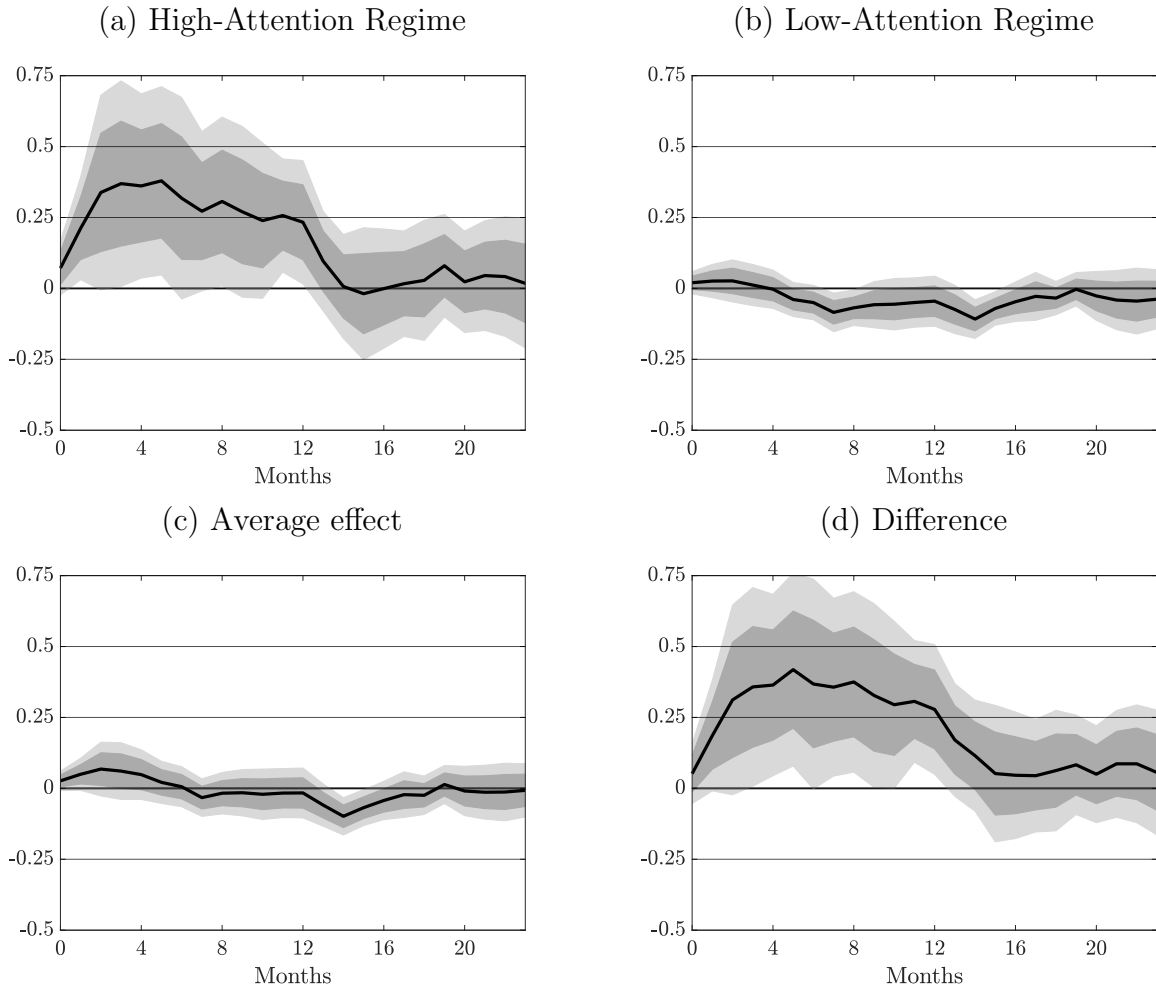
Monetary policy shocks. In figure 13, I show that inflation also responds more strongly in the high-attention regime to monetary policy shocks identified using a high-frequency identification (the shocks are taken from Jarociński and Karadi (2020) and are purged from the information effects of monetary 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. In particular, there are 49 months out of 353 in which quarter-on-quarter CPI inflation was above 4% during that period, so about 10% of the time, whereas in the full sample the economy spends about 30% of the time in the high-attention regime. 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 12: Price level response to the shock targeting inflation

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference
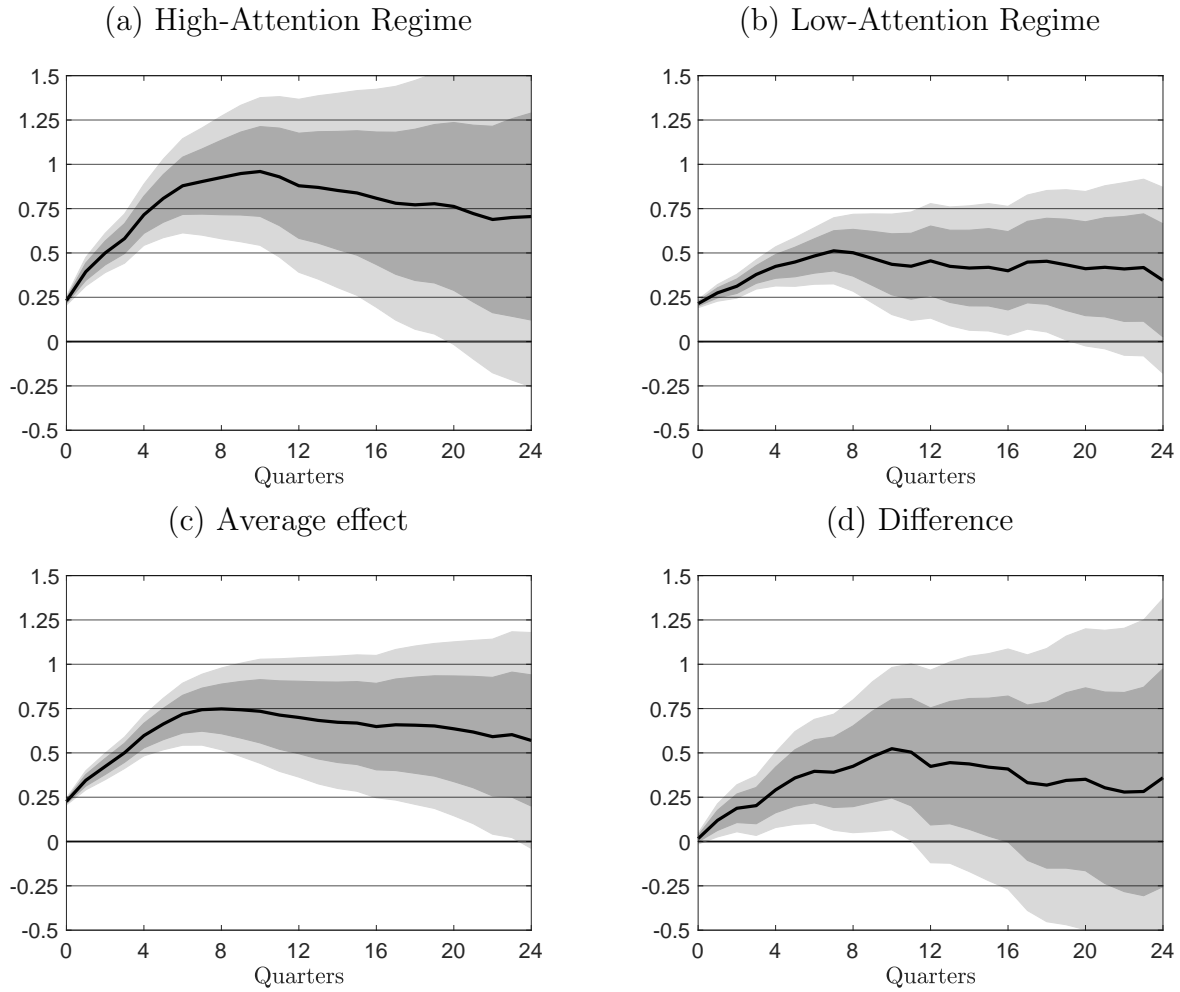
Notes: This figure shows the cumulative price level (using the GDP deflator) response to the shock of Angeletos et al. (2020a) that targets inflation in the high-attention regime (panel (a)), the low-attention regime (panel (b)), on average across regimes (panel (c)), and the difference between the two 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 13: Inflation response to an monetary policy shock

(a) High-Attention Regime

(b) Low-Attention Regime

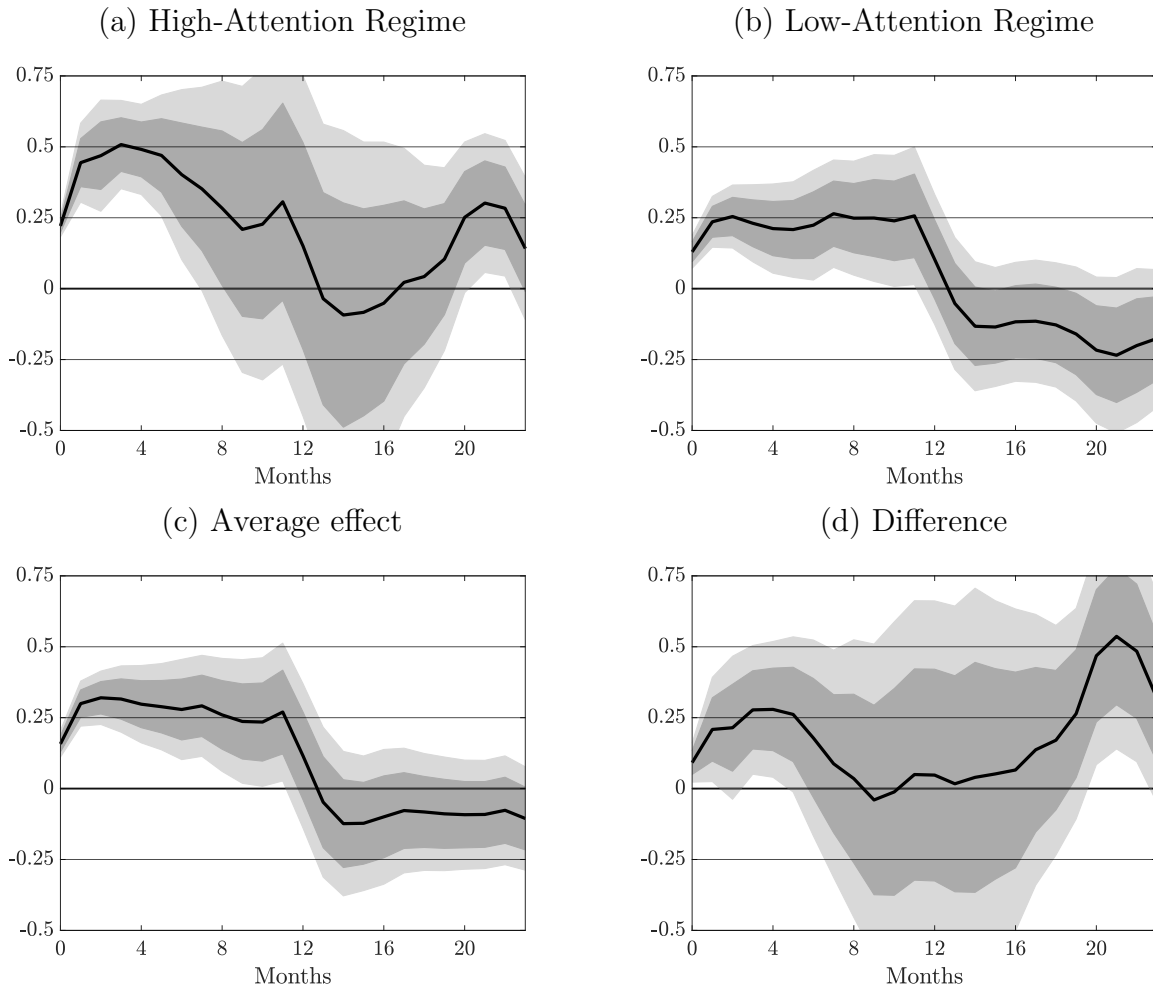(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to a monetary policy shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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

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. As this data is only available since 2004, the high-attention regime largely coincides with the recent inflation surge. I use the same control variables as in my baseline specification and I further include four lags of the number of Google searches as control variables. Figure 14 shows the results. They confirm the baseline results that inflation responds substantially and significantly more to supply shocks when attention to inflation is high.

Figure 14: Inflation response to an oil supply shock with Google data as threshold-defining variable

(a) High-Attention Regime

(b) Low-Attention Regime

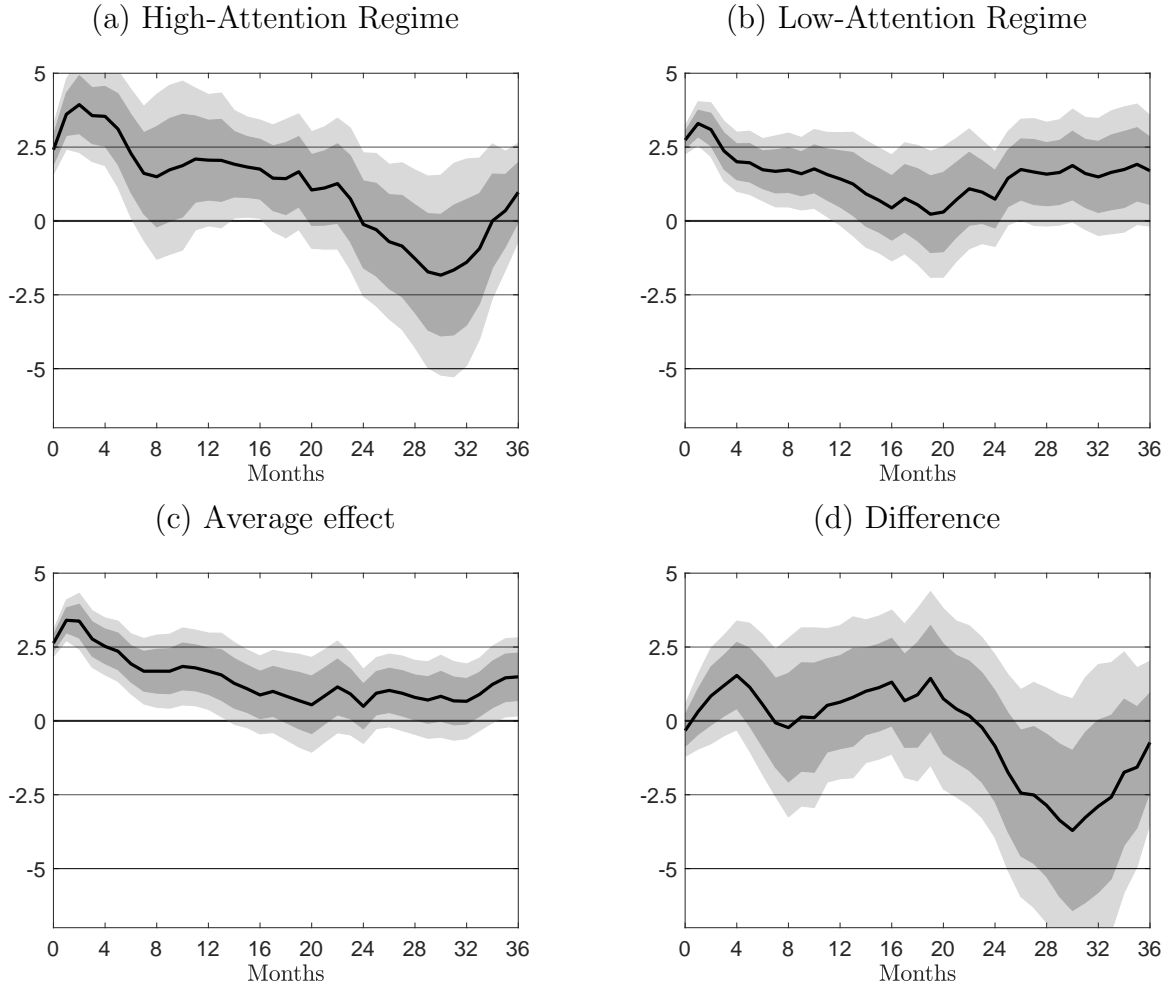(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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).

Oil price response. Figures 15 and 16 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)),

51

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 15: Real oil price response to an oil supply news shock

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the real oil price response to an oil supply news shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 16: Nominal oil price response to an oil supply news shock

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect
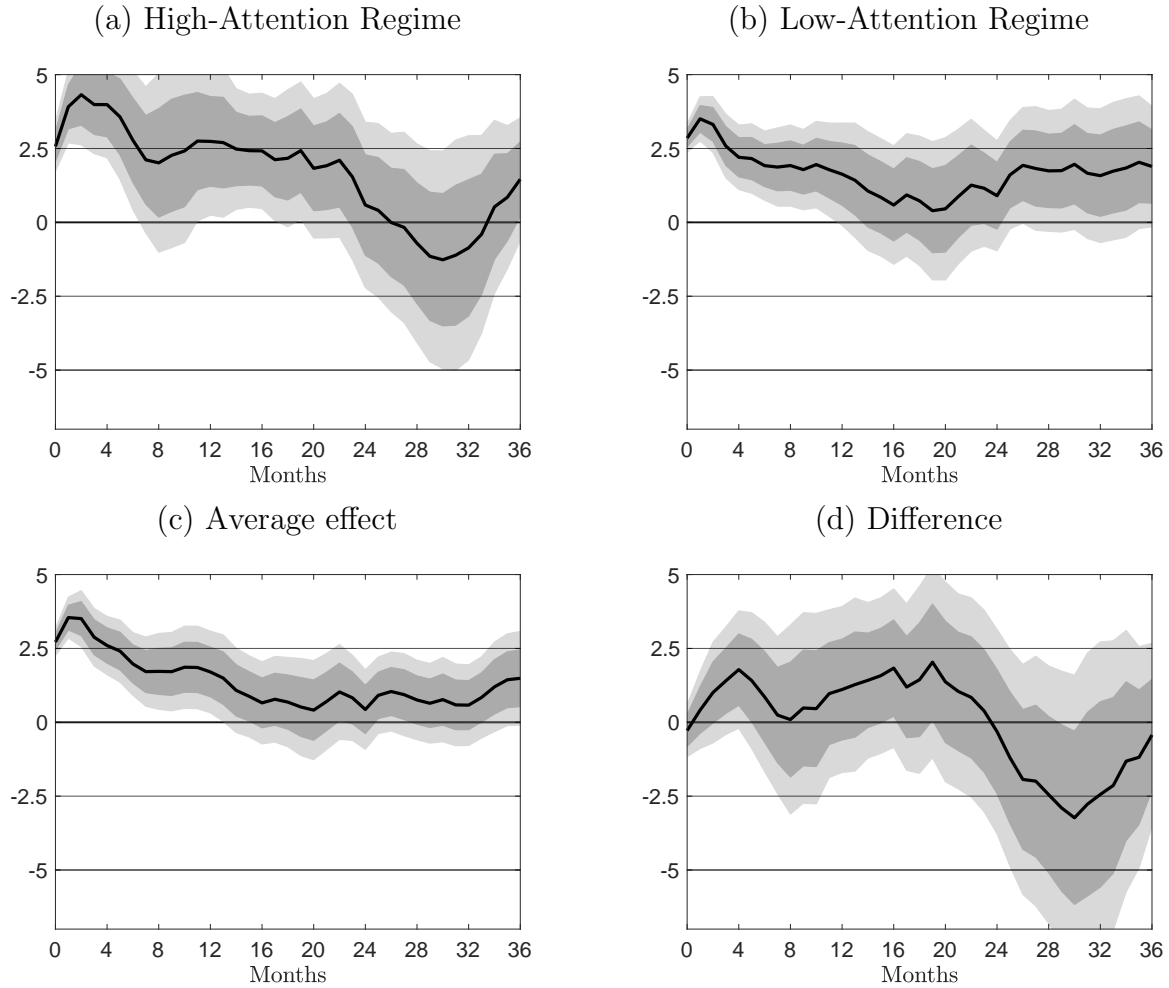
(d) Difference

Notes: This figure shows the nominal oil price response to an oil supply news shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 17 and 18 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 17: Unemployment response to an oil supply news shock
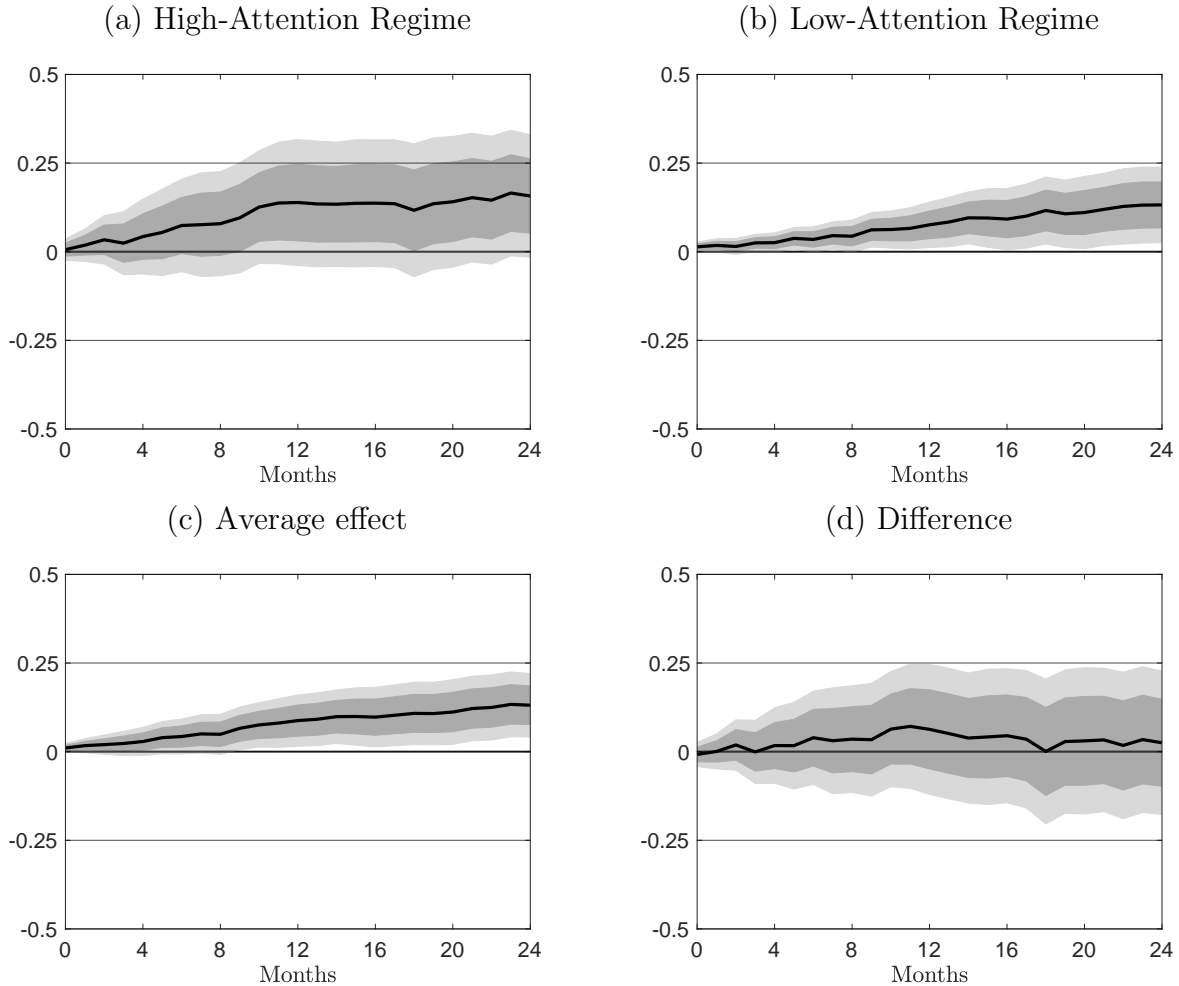
(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the unemployment rate response to an oil supply news shock when excluding the Covid-19 period (2020-2023) in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 18: Unemployment response to an oil supply news shock including Covid

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference
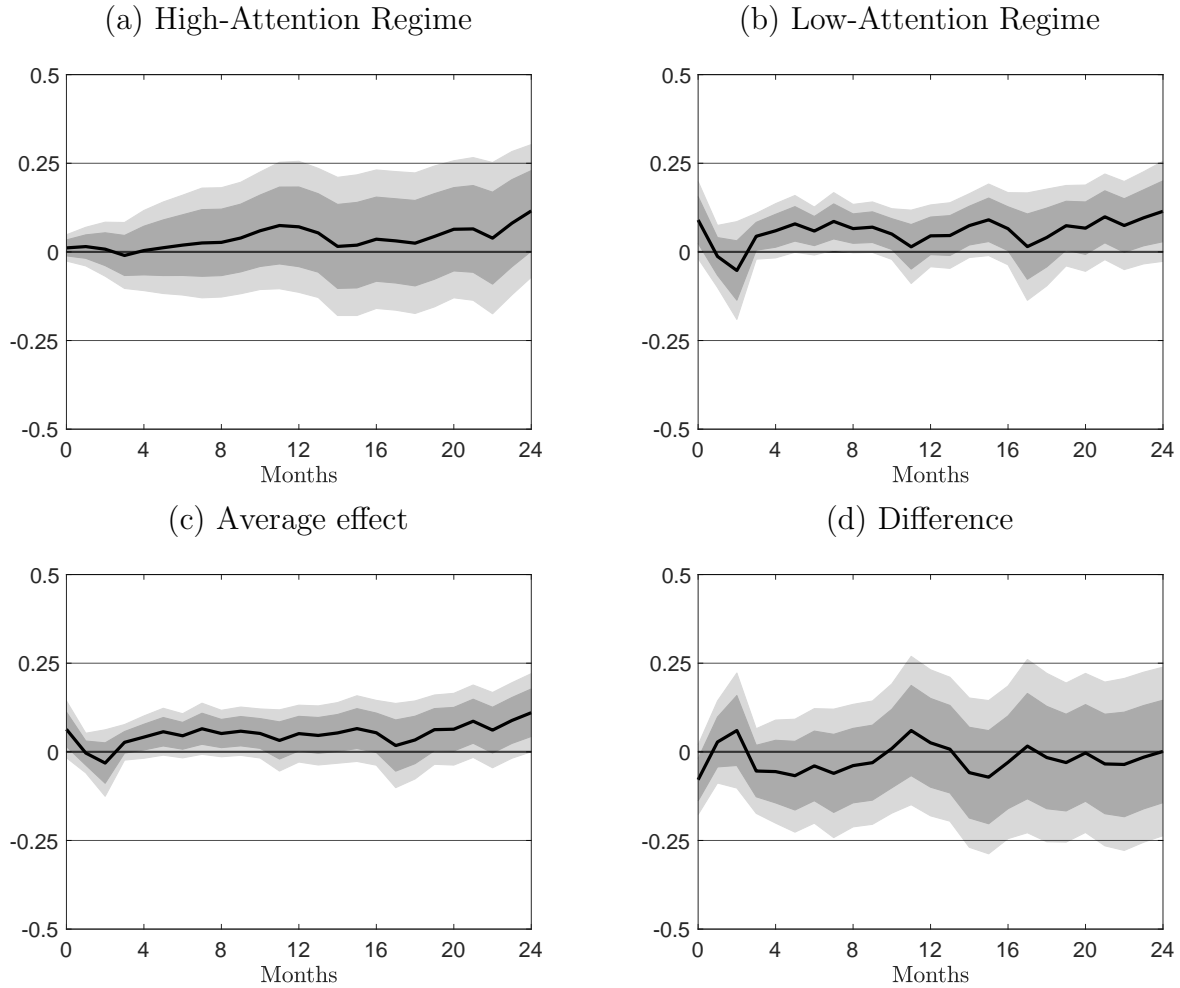
Notes: This figure shows the unemployment rate response to an oil supply news shock when including the Covid-19 period (2020-2023) in the high-attention regime (panel (a)), the low-attention regime (panel (b)), on average across regimes (panel (c)), and the difference between the two regimes (panel (d)). 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 19 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.
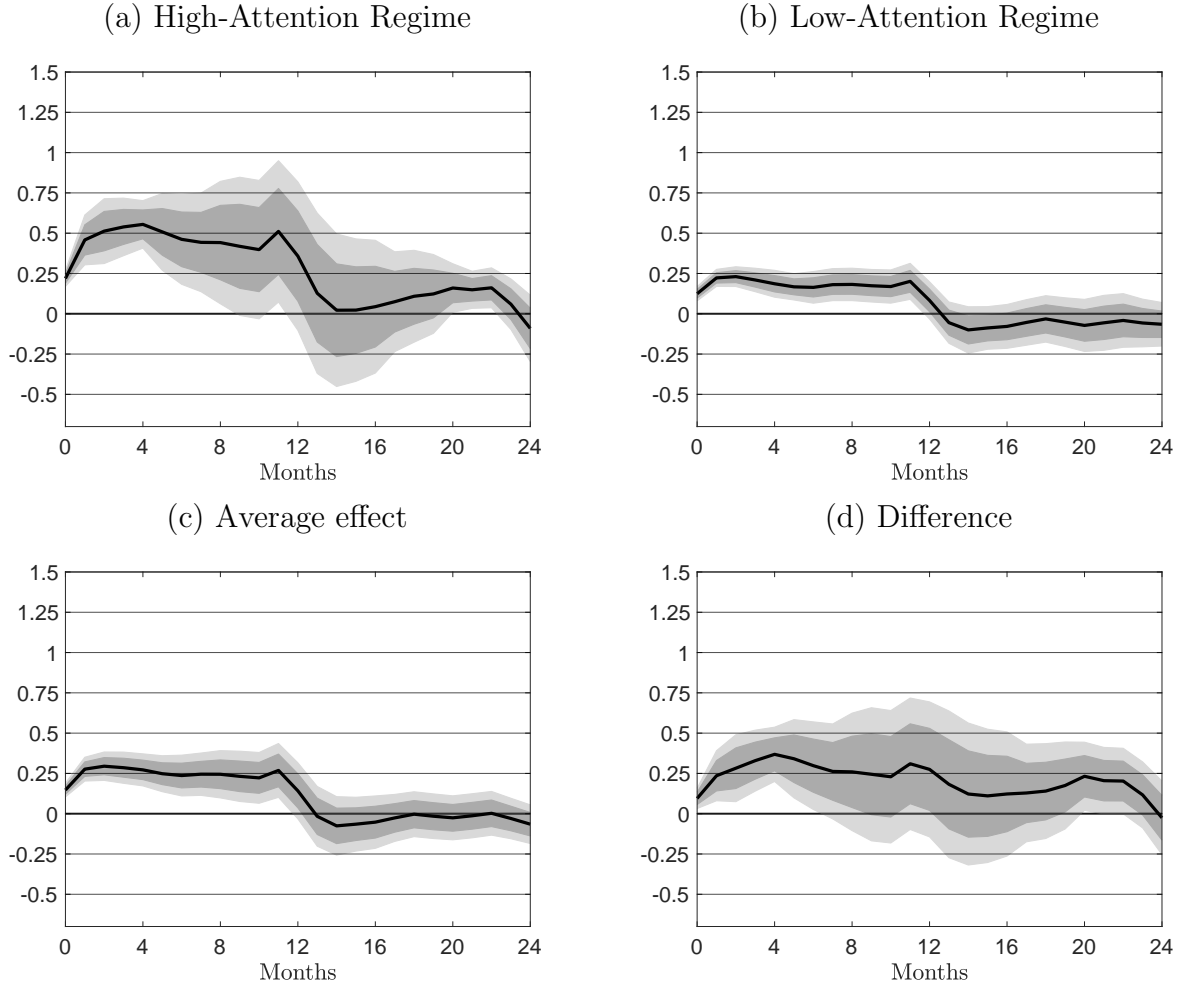
Figure 19: Inflation response to an oil supply news shock excluding Covid
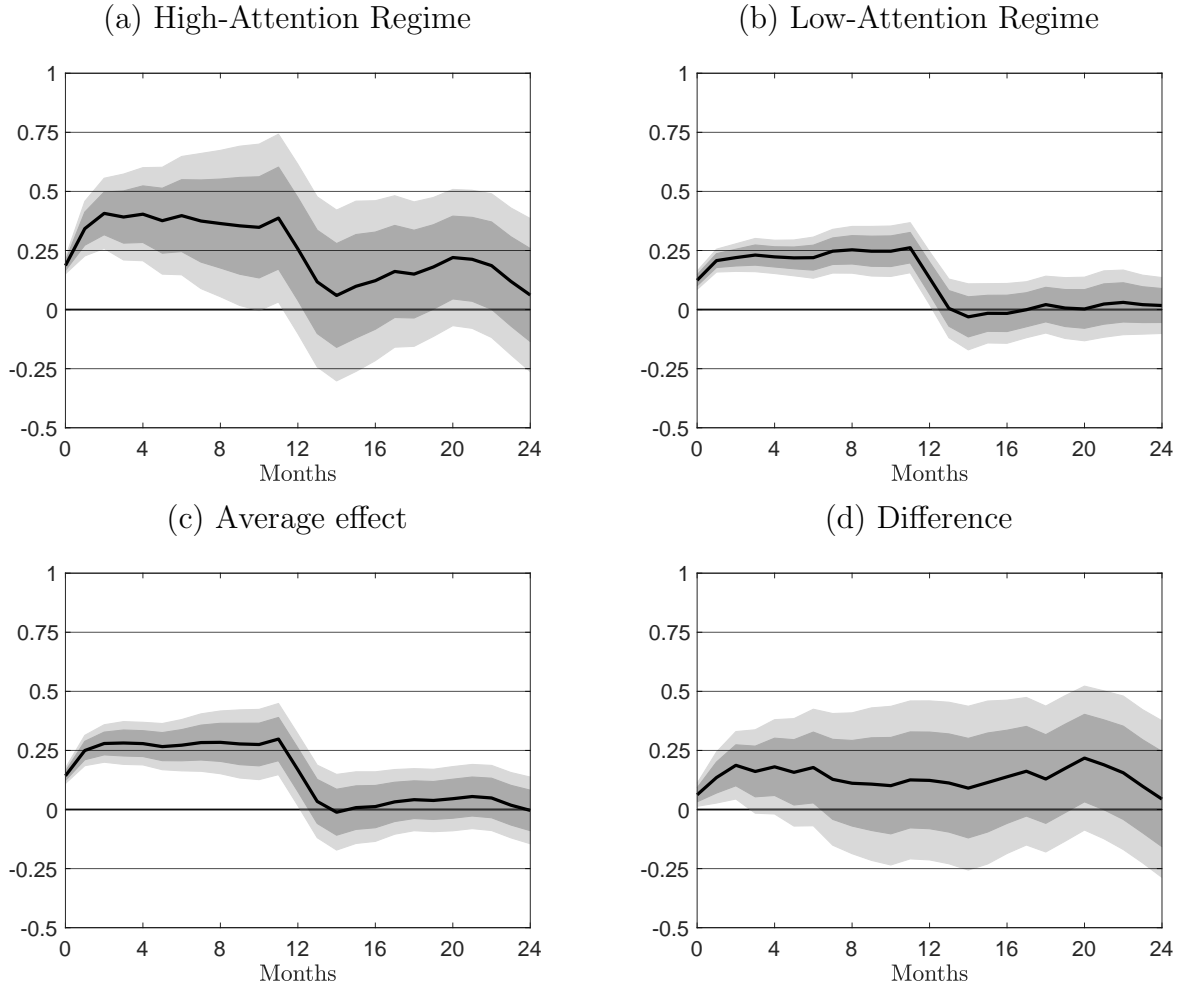
(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock when excluding the years 2020, 2021, 2022 and 2023 in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 20 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.
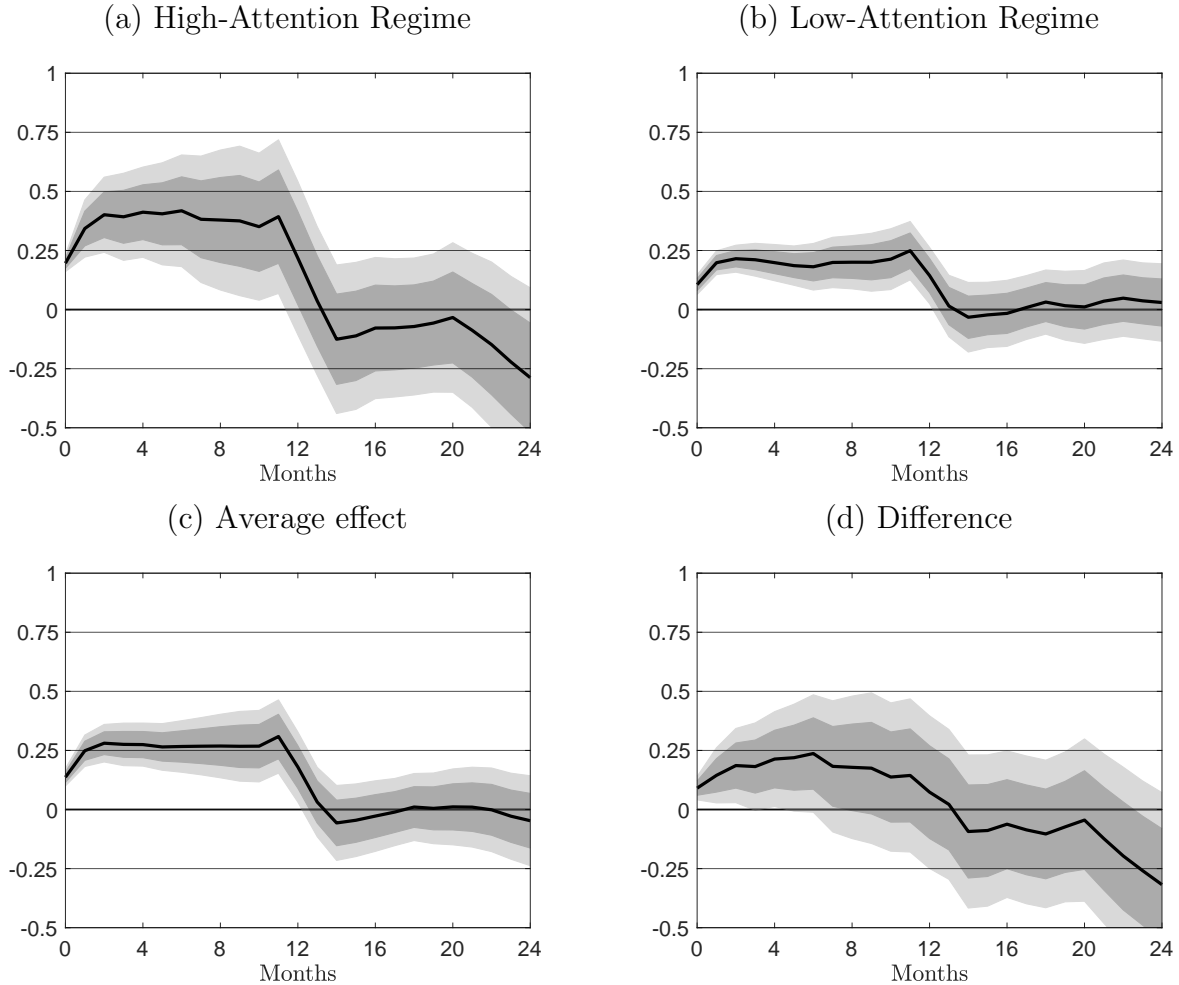
Figure 20: Inflation response to an oil supply news shock after 1990

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock when excluding the years prior to 1990, in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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 21 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.
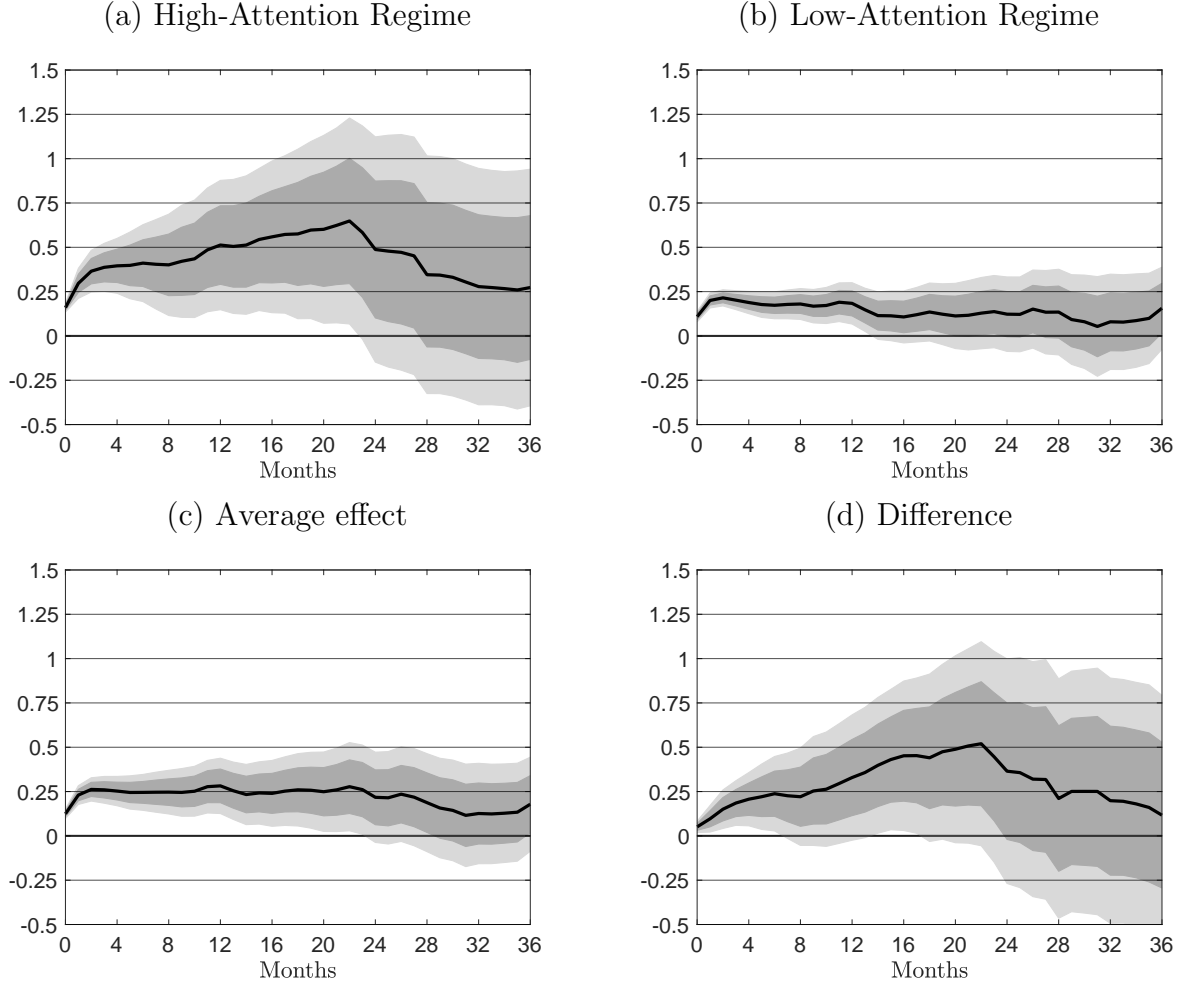
Figure 21: Inflation response to an oil supply news shock (no controls)

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the inflation response to an oil supply news shock when using no control variables in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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. Figure 22 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 22 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 22: Price level response to an oil supply news shock

(a) High-Attention Regime

(b) Low-Attention Regime

(c) Average effect

(d) Difference

Notes: This figure shows the cumulative price level (using the CPI) response to an oil supply news shock in the high-attention regime (panel (a)), the low-attention regime (panel (b)), 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.

Expectations and inflation. Figure 23 shows the response of inflation forecasts to an adverse oil supply news shock using average expectations from the Survey of Consumers (black-solid line) together with the response of actual inflation (blue-solid line). The response of expectations is shifted such that the vertical distance between the two lines captures forecast errors. Consistent with the model, we see that expectations initially undershoot, followed by a delayed overreaction.

Figure 23: Inflation and expected inflation
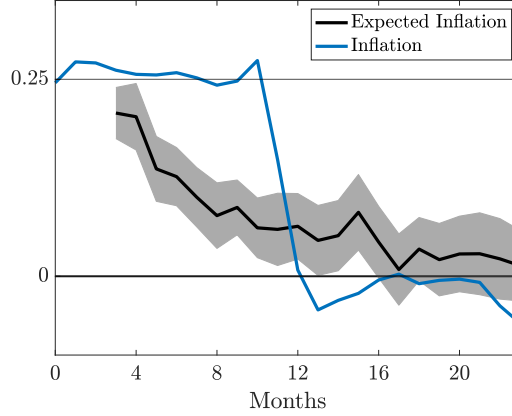
Notes: This figure shows the response of inflation forecasts using average expectations from the Survey of Consumers (black-solid line) together with the response of actual inflation (blue-solid line). The response of expectations is shifted such that the vertical distance between the two lines captures forecast errors. The shaded areas depict the 68% confidence bands, standard errors are robust with respect to serial correlation and heteroskedasticity (Newey and West (1987) with 12 lags).

D Model Derivations and Calibration of γ_x

Derivation of equation (33). Linearizing the Euler equation (18) yields

$$\widehat{c}_t = \tilde{E}_t\widehat{c}_{t+1} - \left(\tilde{i}_t - \tilde{E}_t\pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1})\right). \tag{36}$$

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 $\tilde{E}_t\widehat{z}_{t+1} = E_t\widehat{z}_{t+1} = \rho_z\widehat{z}_t$.

In order to express equation (36) 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 (19), we get that potential output is therefore also constant and equal to

$$Y_t^* = \Xi^{-1}. \tag{37}$$

60

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

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

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

$$r_t^* = \widehat{z}_t - \tilde{E}_t z_{t+1}. \tag{38}$$

Substituting $\widehat{z}_t - \tilde{E}_t z_{t+1}$ with r_t^* in (36) and using that $\widehat{c}_t = \widehat{y}_t = \widehat{x}_t$, since potential output is 0, yields the IS equation (33):

$$\widehat{x}_t = \tilde{E}_t \widehat{c}_{t+1} - \left(\tilde{\imath}_t - \tilde{E}_t \pi_{t+1} - r_t^*\right). \tag{39}$$

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 (39) as

$$\widehat{x}_t = \tilde{E}_t \widehat{x}_{t+1} - \left(\tilde{\imath}_t - \tilde{E}_t \pi_{t+1} - r_t^*\right), \tag{40}$$

which is equation (33) in the main text.

Derivation of equation (32). To derive the New Keynesian Phillips Curve (32), 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

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

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

$$T(\epsilon - 1)\frac{Y}{P}\left[\widehat{T}_t - \epsilon\widehat{p}_t(j) + \widehat{y}_t - (1 - \epsilon)\widehat{p}_t\right] \tag{41}$$

$$= \epsilon mc \frac{Y}{P}\left[(-\epsilon - 1)\widehat{p}_t(j) - (-\epsilon - 1)\widehat{p}_t + \widehat{mc}_t + \widehat{y}_t - \widehat{p}_t\right] \tag{42}$$

$$- \psi\frac{Y}{P}\left(\widehat{y}_t - \widehat{p}_{t-1}(j) - \pi_t(j) - \widehat{y}_t + \widehat{p}_{t-1}(j)\right) \tag{43}$$

$$+ \beta\psi\frac{Y}{P}\tilde{E}_t^j \pi_{t+1}^j. \tag{44}$$

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

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

Given the assumptions discussed in Section 4, it follows that $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t \pi_{t+1}$ and $\widehat{p}_{t-1}(j) = \widehat{p}_{t-1}$ such that

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

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

$$\psi \underbrace{(\widehat{p}_t - \widehat{p}_{t-1})}_{=\pi_t} = \epsilon \left(\widehat{mc}_t - \widehat{T}_t \right) + \beta \psi \tilde{E}_t \pi_{t+1}. \tag{47}$$

Dividing by ψ yields

$$\pi_t = \frac{\epsilon}{\psi} \left(\widehat{mc}_t - \widehat{T}_t \right) + \beta \tilde{E}_t \pi_{t+1}. \tag{48}$$

Using $\widehat{mc}_t = \widehat{y}_t$ 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 $\kappa \equiv \frac{\epsilon}{\psi}$, I arrive at the linearized New Keynesian Phillips Curve under subjective expectations:

$$\pi_t = \kappa \widehat{x}_t + \beta \tilde{E}_t \pi_{t+1} + u_t. \tag{49}$$

D.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. (forthcoming) 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 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 $\mathcal{N}(\mu_t, (\sigma_t)^2)$ 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

$$\sigma_t = \frac{2a}{\Phi^{-1}\left(1 - q_t^U\right) - \Phi^{-1}\left(q_t^D\right)}$$

$$\mu_t = a - \sigma_t \Phi^{-1}\left(1 - q_t^U\right).$$

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.

E Additional Figures and Model Results

E.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 $\phi_\pi > 1$.

The economy starts in the steady state, with $u_0 = 0$, $r_0^* = 0$ and prior expectations 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) curve—derived by plugging in the inflation expectations in the Phillips Curve—in the initial period is given by

$$AS_0: \quad \pi_0 = \frac{\kappa}{1 - \beta\gamma_{\pi,L}}\widehat{x}_0, \tag{50}$$

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

$$AD_0: \quad \pi_0 = -\frac{1}{\phi_\pi - \gamma_{\pi,L}}\widehat{x}_0. \tag{51}$$

Panel (a) in Figure 24 depicts this initial situation graphically.[36] 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

$$AS_1: \quad \pi_1 = \frac{\kappa}{1 - \beta\gamma_{\pi,L}}\widehat{x}_1 + \frac{1}{1 - \beta\gamma_{\pi,L}}u_1$$

$$AD_1: \quad \pi_1 = -\frac{1}{\phi_\pi - \gamma_{\pi,L}}\widehat{x}_1.$$

This situation is shown in panel (b) of Figure 24. 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 that inflation exceeds the threshold.
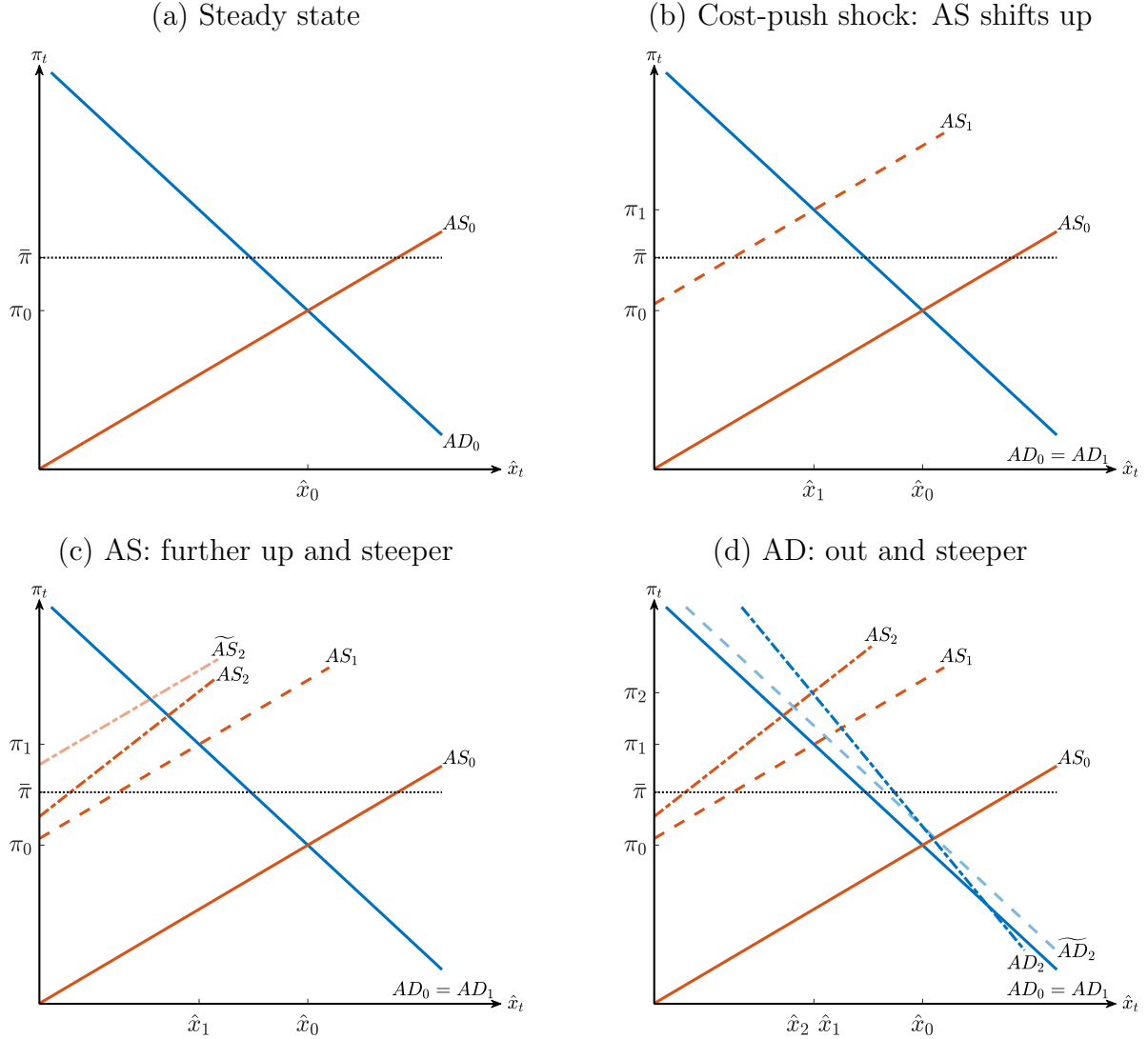
Due to the increase in inflation in period 1, firms enter period 2 with positive prior inflation expectations: $\tilde{E}_1\pi_2 = \gamma_{\pi,L}\pi_1 > 0$. These higher prior expectations together with the still ongoing cost-push shock shift the AS curve further up. This shift is illustrated in panel (c) of Figure 24 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

---

[36]The values I use 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 $u_1 = u_2 = 10$.

Figure 24: An illustrative example

(a) Steady state

(b) Cost-push shock: AS shifts up

(c) AS: further up and steeper

(d) AD: out and steeper

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 $\bar{\bar{\pi}}$ depicts the inflation attention threshold.

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, captured by $\frac{\beta(1-\gamma_{\pi,H})\gamma_{\pi,L}}{1-\beta\gamma_{\pi,H}}\pi_1$, 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 because the increase in attention also steepens the slope of the AS curve. That is, the aggregate supply curve becomes steeper in periods of high attention—a *dynamic* non-linearity. 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 24 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.

It turns out that the higher prior expectations endogenously lead to such a demand increase. This is illustrated by $\widetilde{AD}_2$ in panel (d) of Figure 24, 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

66

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 steep. 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. Thus, the attention threshold offers a potential mechanism for how inflation surges may occur and exhibit self-reinforcing dynamics without large changes in output.

After the shock. 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 increased 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.

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 disinflation occurs more gradually. These results are shown graphically in Figure 25.

Figure 25: Illustrative example: after the shock has died out
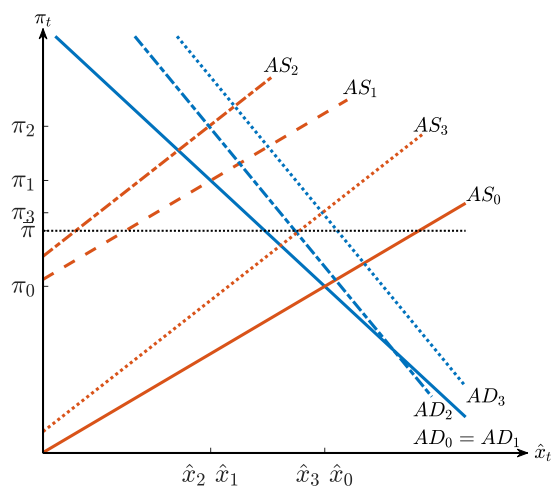

Notes: This figure depicts the third period in which the cost-push shock has died out (after being positive for two consecutive periods). The black-dotted line at $\bar{\pi}$ depicts the inflation attention threshold.

E.2   Additional figures to Section 5.1

Figure 26 shows the inflation response to a cost-push shock for five different models. My model with the attention threshold (labelled "Baseline with threshold" and depicted by the red-dashed line) shows the discussed hump-shaped pattern of the inflation surge and the decrease in the speed of disinflation once inflation falls back below the threshold. As the other four illustrate, these inflation dynamics are unique to the model with the attention threshold. In particular, the models under FIRE (black-solid line) or in which attention remains low (blue-dashed-dotted line) predict that inflation peaks on impact and then comes down relatively quickly. The same is true for the model with a constant attention parameter of 0.24, which corresponds to the estimated attention level in the data if I do not impose a threshold (the inflation response is captured by the blue-dotted line). Additionally, models with FIRE could not account for the observed dynamics of forecast errors. A model variant in which attention to inflation is always high (red-dotted), on the other hand, can produce the initial hump-shaped response but then predicts a very fast decline back to target, i.e., this model would not explain why we see such a long last mile of disinflation. Furthermore, a model in which attention is always high would have predicted a very pronounced deflationary period after the Great Financial crisis. However, once thing that the model with high-attention illustrates, is that the inflation response is hump shaped when attention is high whereas it peaks on impact when the shock hits in times of low attenion, consistent with the empirical findings (panels (a) and (b) in Figure 4).

Figure 26: Inflation dynamics under FIRE or absent the attention change
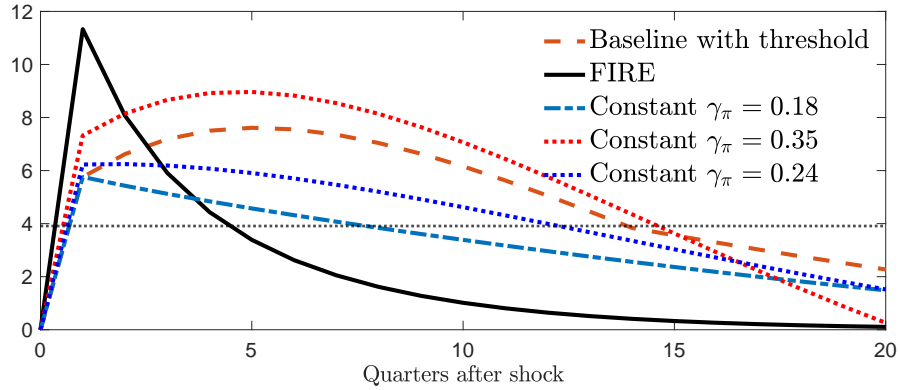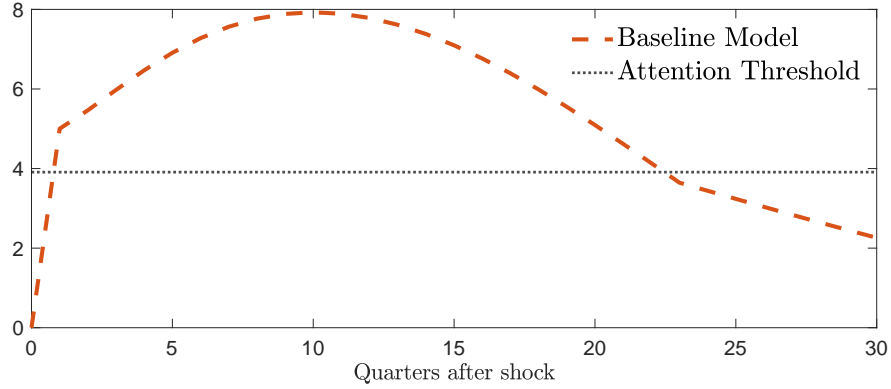

Notes: This figure shows the inflation response to a cost-push shock for different model specifications.

Figure 27 shows the inflation response to a positive demand shock. We see that the discussed inflation dynamics after a relatively large supply shock are similar to the ones after a demand shock: inflation keeps on increasing, fueled by the attention increase, and the speed of disinflation slows substantially once inflation falls back below the threshold and attention therefore decreases. Note, however, that the inflation surge is somewhat longer lived than in response to the supply shock.

Figure 28 shows the inflation response for the case in which attention to the output gap, γ_x, decreases when attention to inflation increases. This may reflect that agents shift their attention from the output gap to inflation when inflation is high. We see that the inflation surge becomes more

Figure 27: Inflation dynamics after a demand shock

Notes: This figure shows the inflation response in the model with the attention threshold in response to a demand shock r_t^* with persistence $\rho_r = 0.55$.

persistent in that case (note that the x-axis now shows 30 periods rather than 20). The reason is that the negative output gap due to the adverse cost-push shock leads to a smaller downward revision of output gap expectations, thus, agents remain relatively more optimistic about the output gap, and hence, consume relatively more, leading to additional upward pressure on inflation.
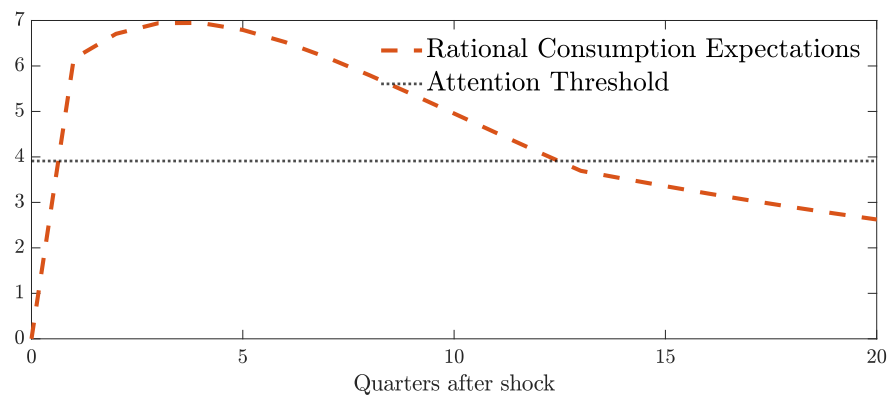
Figure 28: Inflation dynamics with $\gamma_{x,L} = 0.25$ and $\gamma_{x,H} = 0.1$
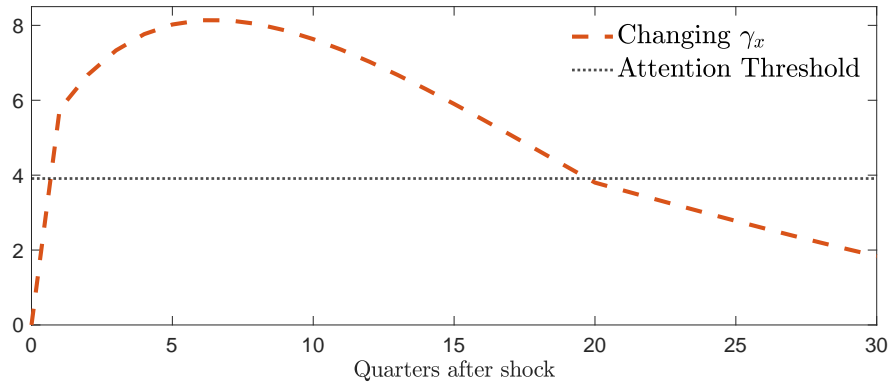
Notes: This figure shows the inflation response to a cost-push shock when attention to the output gap decreases in times of high attention to inflation.

Figure 29 shows the inflation response to a cost-push shock when consumption expectations are formed under full-information rational expectations. We see that the dynamics are qualitatively similar to the baseline model. The endogenous inflation increase is somewhat muted here because the forward-looking consumption expectations do not induce endogenous persistence in the way they do in the baseline model. When increasing the exogenous persistence of the shock slightly, however, the models are very similar (not shown).

Figure 29: Inflation dynamics with rational consumption expectations

Notes: This figure shows the inflation response to a cost-push shock when consumption expectations are formed under full-information rational expectations.