Online Appendix

Inflation - who cares?

Monetary Policy in Times of Low Attention

Oliver Pfäuti*

A A LIMITED ATTENTION MODEL OF INFLATION EXPECTATIONS

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

$$\pi' = \rho_\pi \pi + \nu, \tag{A1}$$

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

$$\pi^{e*} \equiv \rho_\pi \pi. \tag{A2}$$

*Department of Economics, The University of Texas at Austin, pfaeuti.oliver@gmail.com.

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

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

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

$$\max_{f} \int U(\pi^e, \pi) f(\pi^e, \pi) d\pi d\pi^e - C(f) \tag{A3}$$

$$\text{subject to} \int f(\pi^e, \pi) d\pi^e = g(\pi), \text{ for all } \pi, \tag{A4}$$

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

$$C(f) = \lambda I(\pi; \pi^e) = \lambda \left(H(\pi) - E\left[H(\pi|\pi^e)\right]\right), \tag{A5}$$

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

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

$$U(\pi^e, \pi) = -r\left(\rho_\pi \pi - \pi^e\right)^2, \tag{A6}$$

where r measures the stakes of making a mistake.[1,2]

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

$$s = \pi + \varepsilon, \tag{A7}$$

with $\varepsilon \sim i.i.N.(0, \sigma_\varepsilon^2)$.[3] The problem (A3) now reads

$$\max_{\sigma_{\pi|s}^2 \leq \sigma_\pi^2} E_\pi \left[E_s \left[-r\rho_\pi^2 \left(\pi - E[\pi|s]\right)^2\right]\right] - \lambda I(\pi; \pi^e) = \max_{\sigma_{\pi|s}^2 \leq \sigma_\pi^2} \left(-r\rho_\pi^2 \sigma_{\pi|s}^2 - \frac{\lambda}{2} log \frac{\sigma_\pi^2}{\sigma_{\pi|s}^2}\right). \tag{A8}$$

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

$$\pi^e = \rho_\pi \left(1 - \gamma\right) \hat{\pi} + \rho_\pi \gamma s, \tag{A9}$$

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

An equivalent way of writing γ is

$$\gamma = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_\varepsilon^2}. \tag{A10}$$

Now, since the agent *chooses* the level of attention, we can re-formulate (A8) as

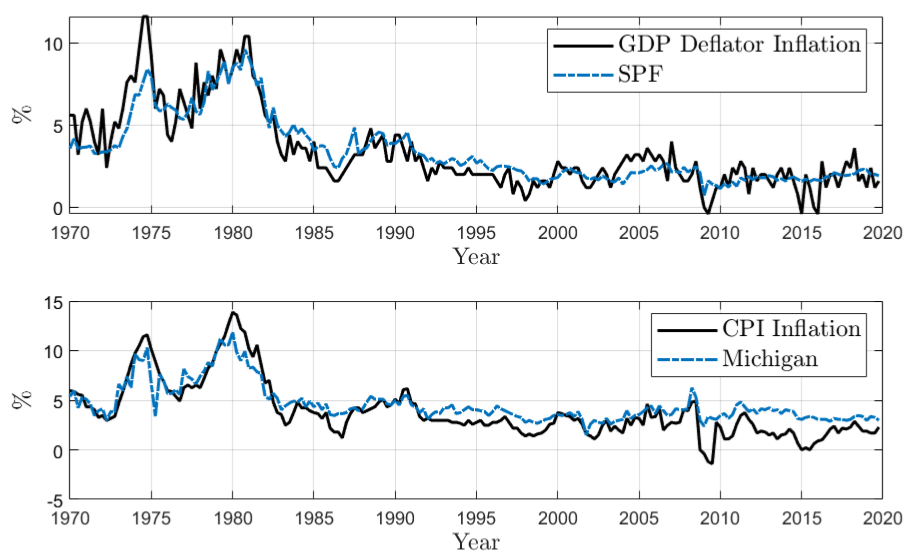
$$\max_{\gamma \in [0,1]} \left(-r\rho_\pi^2 (1 - \gamma) \sigma_\pi^2 - \frac{\lambda}{2} log \frac{1}{1 - \gamma}\right). \tag{A11}$$

Writing the cost of information relative to the stakes, $\tilde{\lambda} \equiv \frac{\lambda}{r}$, and solving the optimization problem (A11) yields the *optimal* level of attention, presented in Lemma 1.

B Appendix to Empirical Results

Figure B1 shows the main time series that are used in the empirical analyses of Section 2. Apart from the apparent decrease in the level and volatility of inflation as well as inflation expectations, we see that expectations became more and more detached from actual inflation. First, consumer expectations seem to be biased on average in the most recent decades, as can be seen in the lower panel. While these expectations closely tracked inflation in the 70s and 80s, this is not the case anymore.[4] Second, professional forecasters' expectations seem to perform quite well on average. In the last twenty years, however, they barely react to actual changes in inflation anymore. Overall, these observations suggest that attention decreased in the last decades.

Figure B1: Inflation and Inflation Expectations



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

Table B1 shows the summary statistics, for the period before and after the 1990s, separately. For professional forecasters, the perceived persistence is higher than the actual one. This is especially the case when the actual persistence is relatively low, as was the case after 1990. Afrouzi et al. (2023) document a similar finding in an experimental setting. This might point towards lower attention since the 1990s. Note, that in the empirical analysis I account for changes in the perceived persistence.

Table B1: Summary statistics

	GDP Deflator Inflation		SPF Expectations	
	1968-1990	1990-2020	1968-1990	1990-2020
Mean (%)	5.44	2.00	5.18	2.16
Std. Dev. (%)	2.43	0.90	1.87	0.64
Persistence	0.84	0.55	0.93	0.92
	CPI Inflation		Consumer Expectations	
	1968-1990	1990-2020	1968-1990	1990-2020
Mean (%)	6.09	2.43	6.00	3.63
Std. Dev. (%)	3.00	1.26	2.17	0.68
Persistence	0.96	0.77	0.85	0.70

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

B.1 Robustness and Additional Evidence

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

Additional Data Sources

In Table B2, I show how attention changed over time for different data sources. The first two columns show the results for the Greenbook forecasts, columns 3-4 for the Livingston Survey, and columns 5-6 and 7-8 are for CPI forecasts from the SPF instead of forecasts about the GDP deflator. As in the main text, I use two different estimators. First, the Blundell-Bond estimator (columns 5-6) and pooled OLS (columns 7-8). All standard errors are robust with respect to heteroskedasticity and serial correlation. We see that the main finding of lower attention in inflation expectations in the period after 1990 compared to the period before is robust to these changes in the data source and/or exact variable.[5]

Table B2: Regression results of equation (8)

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

Note: This table shows the results from regression (8) for different data sources. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Different Sample Splits

Table 1 in the main body of the paper shows that attention to inflation declined by focusing on a sample split in 1990. To show that this is robust to the exact split point, Tables B3 and B4 show that the result holds when splitting the sample in 1985 or 1995, respectively. In fact, the decline in attention is even somewhat more pronounced when splitting the sample in 1985. This is in line with the theoretical prediction of the limited-attention model. Namely, the period between 1985 and 1990 was a period of relatively low and stable inflation compared to the period pre 1985 (see Figure B1), and thus, a period in which the model would predict a relatively low level of attention.

Table B3: Regression results of equation (8), pre 1985 vs. post 1985

	Professional Forecasters				Consumers			
	Blundell Bond		Pooled OLS		Averages		Median	
	< 1985	≥ 1985	< 1985	≥ 1985	< 1985	≥ 1985	< 1985	≥ 1985
$\widehat{\gamma}$	0.75	0.37	0.45	0.25	0.77	0.32	0.50	0.26
s.e.	(0.1247)	(0.0399)	(0.0403)	(0.0338)	(0.1688)	(0.0811)	(0.0955)	(0.0562)
N	1914	3887	1914	3887	64	140	27	140

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

Table B4: Regression results of equation (8), pre 1995 vs. post 1995

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

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

Different Specifications of the BB Estimator

In the baseline estimation, reported in Table 1, I included all potential lags for the Blundell-Bond estimation. To show that the results are robust to this specification, I show in Table B5 that for maximum lag lengths of 20 and 10 periods, the estimated attention parameter $\widehat{\gamma}$ is in all cases higher before 1990 compared to the period after 1990.

Table B5: Different maximum lag lengths

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

Note: This table shows the results from regression (8) for different numbers of lags included in the BB estimation. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Time Fixed Effects

To account for potential changes in trend inflation, I include time-fixed effects in regression (8). To do so, recall that (8) is given by

$$\pi^e_{t+1|t,i} = \beta_i + \beta_1 \pi^e_{t|t-1,i} + \beta_2 \left(\pi_t - \pi^e_{t|t-1,i}\right) + u_{i,t}. \tag{A12}$$

To include time fixed effects, I first compute a period-specific persistence parameter, ρ_π. Note, that in (A12), β_1 measures this persistence. Therefore, I subtract $\widehat{\rho}_\pi \pi^e_{t|t-1,i}$ from both sides and then to directly estimate γ, I further divide both sides by $\widehat{\rho}_\pi$:

$$\frac{\pi^e_{t+1|t,i} - \widehat{\rho}_\pi \pi^e_{t|t-1,i}}{\widehat{\rho}_\pi} = \delta_i + d_t + \gamma \left(\pi_t - \pi^e_{t|t-1,i}\right) + v_{i,t}, \tag{A13}$$

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

The estimated attention levels are 0.75 (s.e. 0.0327) for the period before 1990 and 0.61 (s.e. 0.0295) after 1990 if I use the first-order autocorrelation of expected inflation as my measure of ρ_π. If I use the estimate of β_1 from equation (A13) as my measure of ρ_π, the estimated attention before the 1990s is 0.68 (s.e. 0.0252) and the one after the 1990s is 0.46 (s.e. 0.0242). Thus, we see that the decrease in attention is robust to controlling for time-fixed effects, even though the decline is somewhat muted.

When using the first-order autocorrelation of expected inflation as my measure of ρ_π, estimating equation (9) in this way, delivers a point estimate of 0.07 (s.e. 0.0095) that is statistically significant on all conventional significance levels. The estimate for ζ in regression (10) is 0.30 (s.e. 0.0351), statistically significant on all conventional significance levels. When using $\hat{\beta}_1$ from (A13) as the measure of ρ_π, the point estimate of β in equation (9) is 0.06 (s.e. 0.0074) and the estimate of ζ in (10) is 0.29 (s.e. 0.0306), both statistically significant on all conventional levels of significance. Thus, the positive relationships between attention and volatility, as well as between attention and inflation persistence, are robust to controlling for time fixed effects.

Professional Forecasters in the Aggregate

When estimating attention of professional forecasters' average expectations instead of individual ones, we obtain a value of 0.24 (s.e. 0.0421) for the period before 1990 and of 0.09 (s.e. 0.0432) after 1990. Consistent with the main results, attention substantially decreased in recent decades and is about half after 1990 compared to before.

Estimating regression (9) on aggregate SPF data delivers a coefficient of 0.14 (p-value of 0.000) and the estimate of ζ in regression (10) is 1.1 (p-value of 0.000). Thus, the results reported in the main text are robust.

Joint Regressions

Instead of running regressions (9) and (10) separately, I estimate

$$\widehat{\gamma}_t = \alpha + \beta \widehat{\sigma}_{\pi,t} + \zeta \widehat{\rho}_{\pi,t} + u_t. \tag{A14}$$

Table B6 shows that the results are robust to this change in specification.

Table B6: Attention, inflation volatility and inflation persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.04***	0.05***	0.06***
s.e.	(0.0153)	(0.0128)	(0.0151)
$\widehat{\zeta}$	0.59***	0.65***	0.33***
s.e.	(0.0597)	(0.0499)	(0.0780)
N	165	165	163

Note: This table shows the results of regression (A14). Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Controlling for Average Inflation

A potential confounder in regression (A14) above is the average level of inflation. Thus, I now control for the average level of inflation, computed as the average inflation rate in the respective 10-year window. In particular, I run the following regression

$$\widehat{\gamma}_t = \alpha + \beta\widehat{\sigma}_{\pi,t} + \zeta\widehat{\rho}_{\pi,t} + \omega\widehat{\overline{\pi}}_t + u_t, \tag{A15}$$

where $\widehat{\overline{\pi}}_t$ is the estimated average inflation rate. Table B7 reports the results for the professional forecasters. We see that the volatility and the persistence of inflation are positively related with attention and that these relationships are statistically significant even when controlling for the average level of inflation. The average level of inflation, on the other hand, does not have a positive, statistically-significant, effect on the estimated attention when we control for the volatility and persistence of inflation. These results are consistent with the underlying theoretical model.

Table B7: Controlling for average inflation

	Survey of Professional Forecasters	
Estimator	Blundell-Bond	Pooled OLS
$\widehat{\beta}$	0.10^{***}	0.06^{**}
s.e.	(0.0249)	(0.0173)
$\widehat{\zeta}$	0.61^{***}	0.64^{***}
s.e.	(0.0566)	(0.0503)
$\widehat{\omega}$	-0.03^{***}	0.00
s.e.	(0.0069)	(0.0048)
N	165	165

Note: This table shows the results of regression (A15). Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Quasi-Panel of Consumers

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

Table B8 shows the results. The first two columns report the results for the split point in 1990, and the third and fourth column for the split point in 1995. We see that the estimated attention levels using this quasi panel are similar to the ones obtained using average expectations (Table 1).

Table B8: Regression results of equation (8), quasi-panel

	Survey of Consumers			
	< 1990	≥ 1990	< 1995	≥ 1995
$\widehat{\gamma}$	0.77	0.33	0.70	0.29
s.e.	(0.0934)	(0.0263)	(0.1078)	(0.0289)
N	160	480	240	400

Note: This table shows the results from regression (8), estimated using the Blundell and Bond (1998) estimator, for consumers grouped into four groups, based on their income. The standard errors, reported in parentheses, are robust with respect to heteroskedasticity and serial correlation.

Table B9 shows the results of regressions (9) and (10) (first column), as well as of the joint regression (A14), using this quasi panel of consumers. We see that the results are robust and that there is indeed a significantly positive relation between attention and inflation volatility, as well as between attention and inflation persistence.

Table B9: Attention, inflation volatility and inflation persistence

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

Note: This table shows the results of regressions (9), (10) (first column) and (A14) (second column) using a quasi panel of consumers. The attention parameters have been estimated using the BB-estimator. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Volatility and Persistence of Inflation Expectations

Table B10 shows the results of regressions (9) and (10) using the volatility and persistence of inflation expectations instead of actual inflation as independent variables. Standard errors are robust with respect to heteroskedasticity.

Table B10: Attention, inflation volatility and inflation persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.14***	0.16***	0.13***
s.e.	(0.0153)	(0.0098)	(0.0173)
$\widehat{\zeta}$	1.06***	1.21***	0.22***
s.e.	(0.1272)	(0.0838)	(0.0689)
N	165	165	164

Note: This table shows the results of regressions (9) and (10) using the volatility and persistence of inflation expectations instead of actual inflation as dependent variables. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Window Length

As predicted by the underlying model of optimal information acquisition, I showed that there is indeed a positive relationship between attention to inflation and inflation volatility, as well as between attention and inflation persistence. In the baseline specification, I relied on a rolling-window approach in which every window was 10 years. Tables B11 and B12 show that these results are robust to using different window lengths, namely 5 and 15 years.

Table B11: Attention, inflation volatility and inflation persistence

	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.13***	0.15***	0.13***
s.e.	(0.0456)	(0.0171)	(0.0410)
$\widehat{\zeta}$	0.71	0.52***	0.42***
s.e.	(0.4391)	(0.0458)	(0.1556)
N	185	185	184

Note: This table shows the results of regression (A14) using windows of 5 years each. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Table B12: Attention, inflation volatility and inflation persistence

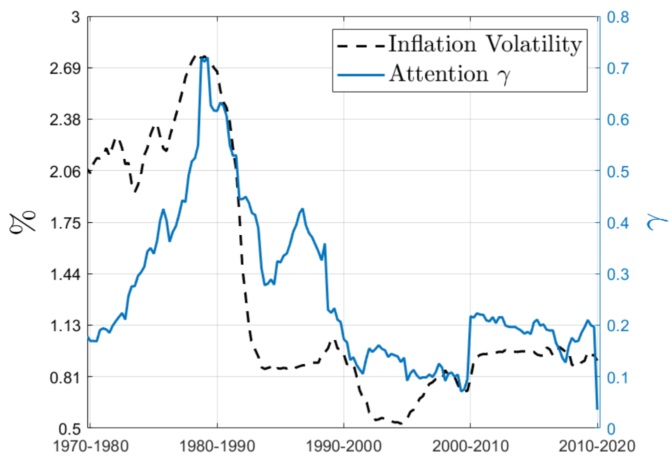
	Survey of Professional Forecasters		Michigan Survey
Estimator	Blundell-Bond	Pooled OLS	OLS
$\widehat{\beta}$	0.01	0.01	0.07***
s.e.	(0.0124)	(0.0115)	(0.0136)
$\widehat{\zeta}$	0.90***	1.00***	0.43***
s.e.	(0.0603)	(0.0552)	(0.0707)
N	145	145	144

Note: This table shows the results of regression (A14) using windows of 15 years each. Standard errors are robust with respect to heteroskedasticity. *** : p-value < 0.01, ** : p-value < 0.05, * : p-value < 0.1.

Attention over Time

Figure B2 shows the estimated attention levels, γ, (blue-solid line) from the SPF consensus forecasts, together with the volatility of GDP deflator inflation (black-dashed line). We clearly see the aforementioned decrease in attention over time, as well as the positive correlation of attention and inflation volatility.

Figure B2: Attention and Inflation Volatility over Time



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

AR(2) Beliefs

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

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

$$\pi_t = \phi_1 \pi_{t-1} + \phi_2 \pi_{t-2} + \nu_t, \tag{A16}$$

and that they receive signals of the form $s_t = \pi_t + \epsilon_t$, and the disturbances ν and ϵ are i.i.d. zero-mean normally-distributed random variables with time-invariant volatilities.[6] Following the arguments in Hamilton (1994), the steady state Kalman filter then yields:

$$\begin{pmatrix} \pi^e_{t+1|t} \\ \pi^e_{t|t} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \pi^e_{t|t-1} \\ \pi^e_{t-1|t-1} \end{pmatrix} + \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} \left(\pi_t + \epsilon_t - \pi^e_{t|t-1}\right). \tag{A17}$$

The second equation, shifted one period backwards, yields

$$\pi^e_{t-1|t-1} = \pi^e_{t-1|t-2} + k_2 \left(\pi_{t-1} + \epsilon_{t-1} - \pi^e_{t-1|t-2}\right), \tag{A18}$$

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

$$\pi^e_{t+1|t} = \phi_1 \pi^e_{t|t-1} + \phi_2 \pi^e_{t-1|t-2} + k_1 \left(\pi_t - \pi^e_{t|t-1}\right) + \underbrace{K_2}_{=\phi_2 k_2} \left(\pi_{t-1} - \pi^e_{t-1|t-2}\right) + u_t, \tag{A19}$$

where k_1 and k_2 are the coefficients in the Kalman gain matrix (denoted by K in Hamilton (1994)), and $u_t = k_1 \epsilon_t + K_2 \epsilon_{t-1}$. I consider household average and median expectations when estimating regression (A19), and I include an intercept. To account for serial correlation in the error term, I apply the Newey-West estimator using four lags (Newey and West (1987)). Table B13 shows the results.

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

Table B13: AR(2) perceived law of motion

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

Note: This table shows the results of regression (A19) for household average and median expectations. Standard errors are robust with respect to heteroskedasticity and serial correlation (Newey-West with 4 lags).

Furthermore, the estimates in Table B13 show that none of the estimated coefficients that arise due to the AR(2) assumption—the estimates for K_2 and ϕ_2—are statistically significant at the 1% level (often, not even at the 5% or 10% level). When estimating attention in exactly the same way as in Section 2, i.e., computing attention as $\frac{k_1}{\phi_1}$, I obtain estimates that are very close to the ones in the main text where I ignore the effects arising from the second lag of inflation in the perceived law of motion. For average expectations, I estimate for the period before 1990 an attention parameter of 0.78 (it was 0.75 in the main text), and for the period after 1990 a value of 0.34 (0.31 in the main text). For median expectations, I obtain attention estimates of 0.41 for the period before 1990 (it is 0.43 in the main text) and 0.255 for the period after 1990 (0.24 in the main text). These findings give empirical support to the assumption that the perceived law of motion for inflation follows an AR(1).

B.2 Other Measures of Attention

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

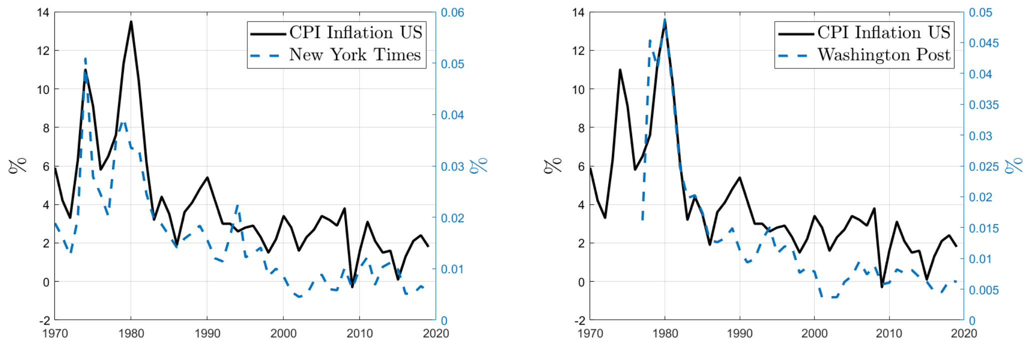
News Coverage of Inflation

Figure B3 shows the relative frequency of the word "inflation" among all words in two major U.S. newspapers (blue-dashed lines), the New York Times (left panel) and the Washington Post (right panel), together with the annual U.S. CPI inflation (black-solid lines). It is evident that news coverage is higher in times of high and volatile inflation as was the case during the 1970s and early 1980s. Moreover, the figure suggests that the public's attention to inflation—proxied here by news coverage—has not always been as low as in recent years, but declined over time.

In Figure B4, we see that a similar picture emerges when looking at the coverage of "inflation" in books, according to *Google Books Ngram Viewer*. In the left panel, we see that "inflation" is covered more frequently in English books written in times of high inflation. But this is not simply a U.S. phenomenon. To see this, I show the same statistic for books written in Spanish for the word "inflación". To contrast this with inflation, the black solid line shows the average inflation (in logs) of the four largest Spanish-speaking countries, weighted by their 2020 population size. These are Argentina, Colombia, Mexico and Spain. Again, we observe that attention to inflation—measured by book coverage—is higher in times of high and volatile inflation.

Figure B3: News Coverage of Inflation

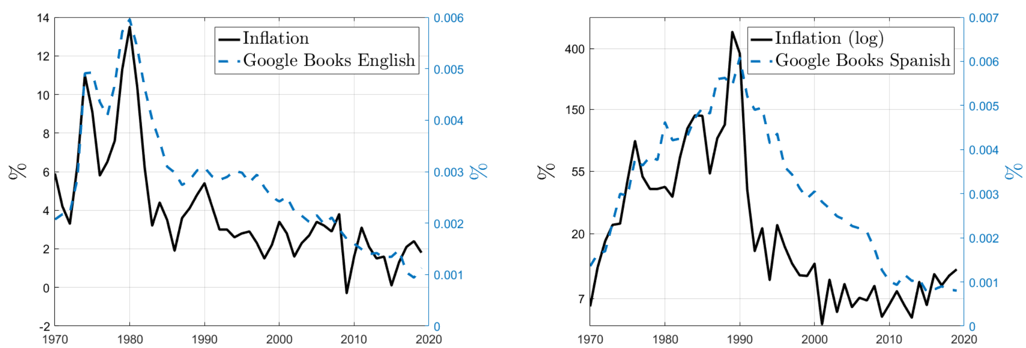
(a) New York Times (b) Washington Post



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

Figure B4: Book Coverage of Inflation

(a) Google Books English (b) Google Books Spanish



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

Households answering "I don't know"

Another potential measure of people's inattention is to see how the share of survey respondents that answer the question about their inflation expectations with "I don't know" changes over time. The Michigan survey provides these shares. Following a rolling-windows approach, I compute for each 10-year window the estimated attention parameter $\widehat{\gamma}_t$ as well as the average share of households within these 10 years that say "I don't know". When using average expectations to estimate γ, I find that the two are strongly negatively correlated with a correlation coefficient of -0.4. When regressing the share of "don't know" respondents on the estimated attention parameter, I obtain a regression coefficient of -7.23 (p-value of 0.010). When controlling for the window-specific inflation volatility, autocorrelation and average inflation rate, the regression coefficient is -5.64 with a p-value of 0.016. When using median expectations to compute γ, the results are even slightly stronger. The raw correlation is -0.47, the regression coefficient without controls is -4.29 with a p-value of 0.000, and when adding controls it equals -3.17 with a p-value of 0.031. These results indicate that inattention (as measured by the share of respondents answering "I don't know") is lower in times my measure of attention, γ, is higher. Thus, these findings support the view that my measure of attention indeed captures people's attention to inflation.

Accuracy of Nowcasts

In a setup in which the agent cannot distinguish between a trend and a cyclical component of inflation with time-varying volatilities of these two components, if the trend component's contribution to overall inflation increases, the agent's forecast would become more responsive to current inflation, too, similarly to an increase in the attention parameter γ. To differentiate these two models, I therefore now also consider *nowcasts* of inflation and their accuracy. The optimal attention choice problem presented in Section 2 says that more attentive agents receive more precise signals about current inflation and should therefore make smaller nowcast errors in times of high attention. This prediction is exclusive to the proposed model of attention and does not apply to the alternative model of the trend and cycle component of inflation. Using the nowcasts from the Survey of Professional Forecasters, I now test this prediction of the model. To do this, I take the absolute value of forecast errors of current inflation or the squared forecast errors as my two measures of the accuracy of the forecasters' nowcasts. I then compute the average across all forecasters and estimate a time series of these average forecast errors using a rolling-windows approach where each window is 10 years long. Similarly, I estimate the window-specific volatility and persistence of perceived inflation. Consistent with my theory of attention, I find strong negative correlations between inflation volatility and forecast errors, as well as between inflation persistence and forecast errors. This holds for both measures of forecast errors, i.e., for the absolute values and the squared values of forecast errors. The correlations are indeed quite strong. For the squared forecast errors, I find a correlation with inflation persistence of -0.59 and with inflation volatility of -0.26. For the absolute values of forecast errors, the correlation with persistence is -0.65 and with inflation volatility -0.41. These results are consistent with the recent findings in Weber et al. (2023) who find that households that report to pay more attention to inflation have inflation expectations that are much closer to the actual level of inflation.

C Model Details and Derivations

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

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

$$\tilde{E}_0 \sum_{t=0}^{\infty} \beta^t Z_t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \Psi \frac{H_t^{1+\nu}}{1+\nu} \right], \tag{A20}$$

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

Households maximize their lifetime utility subject to the flow budget constraints

$$C_t + B_t = w_t H_t + \frac{1+i_{t-1}}{1+\pi_t} B_{t-1} + \frac{T_t^H}{P_t}, \quad \text{for all } t, \tag{A21}$$

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

Maximizing (A20) subject to (A21) yields the Euler equation

$$Z_t C_t^{-\sigma} = \beta(1+i_t)\tilde{E}_t\left[Z_{t+1}C_{t+1}^{-\sigma}\frac{1}{1+\pi_{t+1}}\right], \tag{A22}$$

and the labor-leisure condition

$$w_t C_t^{-\sigma} = \Psi H_t^{\nu}. \tag{A23}$$

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

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}}dj\right)^{\frac{\epsilon}{\epsilon-1}}, \tag{A24}$$

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:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} Y_t. \tag{A25}$$

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

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon}dj\right)^{\frac{1}{1-\epsilon}}. \tag{A26}$$

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

$$Y_t(j) = H_t(j). \tag{A27}$$

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

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

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

$$(1-\tau_t)P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}\frac{Y_t}{P_t} - w_t H_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 Y_t + t_t^F(j), \tag{A28}$$

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

$$Profits_0(j) = \tilde{E}_0^j \sum_{t=0}^{\infty} D_{0,t}\left[(1-\tau_t)P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}\frac{Y_t}{P_t} - mc_t H_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 Y_t + t_t^F(j)\right], \tag{A29}$$

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

$$T_t(\epsilon-1)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} = \epsilon mc_t\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon-1} - \psi\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)\frac{P_t}{P_{t-1}(j)} \tag{A30}$$

$$+\beta\psi\tilde{E}_t^j\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\left(\frac{P_{t+1}(j)}{P_t(j)}-1\right)\frac{P_{t+1}(j)}{P_t(j)}\frac{P_t}{P_t(j)}\frac{Y_{t+1}}{Y_t}\right], \tag{A31}$$

where $T_t \equiv 1-\tau_t$.

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

The real government budget constraint is given by

$$B_t = B_{t-1}\frac{1+i_{t-1}}{\Pi_t} + \frac{T_t^H}{P_t} - \tau Y_t + t_t^f. \tag{A32}$$

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

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

$$C = \left(1 - \frac{\psi}{2}(\bar{\Pi}-1)^2\right) Y. \tag{A33}$$

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

$$mc = \frac{(1-\tau)(\epsilon-1)}{\epsilon} + \psi\bar{\Pi}(\bar{\Pi}-1)\frac{1}{\epsilon}(1-\beta), \tag{A34}$$

which implies that the steady state subsidy is equal to

$$T = 1 - \tau = \frac{\epsilon - \psi(\bar{\Pi}-1)\bar{\Pi}(1-\beta)}{\epsilon - 1}. \tag{A35}$$

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

$$Y = \left(\frac{1}{\Psi\left(1 - \frac{\psi}{2}(\bar{\Pi}-1)^2\right)^{\sigma}}\right)^{\frac{1}{\nu+\sigma}}. \tag{A36}$$

Linearization. Linearizing the Euler equation (A22) yields

$$\widehat{c}_t = \tilde{E}_t\widehat{c}_{t+1} - \varphi\left(\tilde{i}_t - \tilde{E}_t\pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1})\right), \tag{A37}$$

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

$$\widehat{y}_t = \widehat{c}_t + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2}\pi_t, \tag{A38}$$

where π_t now denotes inflation in deviations from its steady state value. Plugging this into (A37), we get

$$\widehat{y}_t = \tilde{E}_t\widehat{y}_{t+1} + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2}\left[\pi_t - \tilde{E}_t\pi_{t+1}\right] - \varphi\left(\tilde{i}_t - \tilde{E}_t\pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1})\right). \tag{A39}$$

In order to express this in terms of the output gap, rather than output, we have to solve for the efficient output that prevails in the economy absent price rigidities (denoted by a "$*$"). From the production function, we have $Y_t^* = H_t^*$. The real wage is constant $w_t^* = 1$. From the labor-leisure equation (A23), we get that potential output is therefore also constant and equal to

$$Y_t^* = \Psi^{-\frac{1}{\nu+\sigma}}. \tag{A40}$$

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

$$0 = -\varphi\left(r_t - (\widehat{z}_t - \tilde{E}_t z_{t+1})\right). \tag{A41}$$

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

$$r_t^n = \widehat{z}_t - \tilde{E}_t z_{t+1}. \tag{A42}$$

Substituting $\widehat{z}_t - \tilde{E}_t z_{t+1}$ with r_t^n in (A37) and using that $\widehat{y}_t = y_t^{gap}$, since potential output in deviations from steady state is 0, yields the aggregate IS equation

$$y_t^{gap} = \tilde{E}_t y_{t+1}^{gap} + \frac{\psi(\bar{\Pi} - 1)\bar{\Pi}}{1 - \frac{\psi}{2}(\bar{\Pi} - 1)^2}\left[\pi_t - \tilde{E}_t\pi_{t+1}\right] - \varphi\left(i_t - \tilde{E}_t\pi_{t+1} - (\widehat{z}_t - \tilde{E}_t z_{t+1})\right). \tag{A43}$$

I assume for the most part of the analysis that output gap expectations are rational, $\tilde{E}_t y_{t+1}^{gap} = E_t y_{t+1}^{gap}$, and that households believe that inflation follows an AR(1) process and that they receive signals of the form $s_t = \pi_t + \varepsilon_t$ with normally distributed noise ε_t (see Section A for details). For tractability, I abstract from noise shocks and therefore assume that $\varepsilon_t = 0$ for all t but that the household behaves as if there was noise. This then gives rise to the law of motion for inflation expectations stated in equation (13). Taking everything together, we can therefore write

$$y_t^{gap} = E_t y_{t+1}^{gap} + \frac{\psi(\bar{\Pi}-1)\bar{\Pi}}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}\left[\pi_t - \pi_{t+1|t}^e\right] - \varphi\left(i_t - \pi_{t+1|t}^e - r_t^n\right). \tag{A44}$$

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

$$y_t^{gap} = E_t y_{t+1}^{gap} - \varphi\left(i_t - \pi_{t+1|t}^e - r_t^n\right), \tag{A45}$$

as stated in equation (12).

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

$$\underbrace{T_t(\epsilon-1)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon}}_{I} = \underbrace{\epsilon mc_t\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon-1}}_{II} \underbrace{-\psi\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)\frac{P_t}{P_{t-1}(j)}}_{III} \tag{A46}$$

$$+\underbrace{\beta\psi\tilde{E}_t^j\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\sigma}\left(\frac{P_{t+1}(j)}{P_t(j)}-1\right)\frac{P_{t+1}(j)}{P_t(j)}\frac{P_t}{P_t(j)}\frac{Y_{t+1}}{Y_t}\right]}_{IV} \tag{A47}$$

The linearization of the terms I to IV yields

$$I: \qquad (\epsilon-1)T\widehat{T}_t - \epsilon(\epsilon-1)T\widehat{p}_t^j + (\epsilon-1)\epsilon T\widehat{p}_t \tag{A48}$$

$$II: \qquad \epsilon(\widehat{mc}_t - (1+\epsilon)\widehat{p}_t^j + (1+\epsilon)\widehat{p}_t) \tag{A49}$$

$$III: \qquad -\psi\bar{\Pi}(\bar{\Pi}-1)\widehat{p}_t - \psi\bar{\Pi}^2\widehat{p}_t^j + \psi\bar{\Pi}(2\bar{\Pi}-1)\widehat{p}_{t-1}^j \tag{A50}$$

$$IV: \quad \beta\psi\bar{\Pi}^2 \tilde{E}_t^j \pi_{t+1}^j + \beta\psi\bar{\Pi}(\bar{\Pi}-1)\Bigg[(1-\sigma)\tilde{E}_t^j \widehat{y}_{t+1} + (\sigma-1)\widehat{y}_t \tag{A51}$$

$$+ \frac{\sigma\psi(\bar{\Pi}-1)}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}\left(\tilde{E}_t^j \pi_{t+1} - \pi_t\right) + \tilde{E}_t^j \pi_{t+1}^j + \widehat{p}_t - \widehat{p}_t^j\Bigg], \tag{A52}$$

where I used the linearized resource constraint to arrive at the expression IV. When trend inflation is zero, $\bar{\Pi} = 1$, the expression IV becomes simply $\beta\psi\tilde{E}_t^j \pi_{t+1}^j$. Since, I focus on the case $\sigma = 1$, the terms relating to output in expression IV drop out. Thus, the only reason why prices may differ across firms j is due to different forecasts of future inflation (either of aggregate or of firm-specific inflation). Following the assumption in Adam and Padula (2011), I assume that these forecasts are provided by forecasters that are different from the price setting managers. Given that there are no idiosyncratic shocks, I assume that the forecaster of firm j expects firm-specific inflation to be equal to aggregate inflation, $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t^j \pi_{t+1}$. These forecasters then form their inflation expectations in the same way as households and as detailed in the limited attention problem in Section 2. All forecasters receive the same signal, and I assume that this signal is perfectly accurate but forecasters perceive the signal as being noisy. Therefore all forecasters form their expectations equally. Note, that a weaker assumption would be sufficient to arrive at the following results: namely, that all forecasters receive the same signal.[7] Thus, $\tilde{E}_t^j \pi_{t+1}^j = \tilde{E}_t^j \pi_{t+1} = \tilde{E}_t \pi_{t+1}$. Therefore, all price setters set the same price, $\widehat{p}_t^j = \widehat{p}_t$. Thus, the belief that firm-specific inflation coincides with aggregate inflation is satisfied in equilibrium, which confirms the forecasters' belief and she therefore does not have an incentive to update that belief.

Using this, and

$$\widehat{mc}_t = (1+\nu)\widehat{y}_t - \frac{\psi(\bar{\Pi}-1)\bar{\Pi}}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}\pi_t, \tag{A53}$$

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

$$\pi_t = \zeta \left[\frac{\epsilon(1+\nu)}{\psi} \widehat{y}_t^{gap} + \beta \bar{\Pi}^2 \pi^e_{t+1|t} + u_t + \Xi \pi^e_{t+1|t} \right] \tag{A54}$$

where

$$\zeta \equiv \frac{1}{\bar{\Pi}(2\bar{\Pi}-1) + \frac{\beta\psi\bar{\Pi}(\bar{\Pi}-1)^2}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2} + \frac{\epsilon\bar{\Pi}(\bar{\Pi}-1)}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}} \tag{A55}$$

$$\Xi \equiv \beta\bar{\Pi}(\bar{\Pi}-1)\left[1 + \frac{\psi(\bar{\Pi}-1)}{1-\frac{\psi}{2}(\bar{\Pi}-1)^2}\right], \tag{A56}$$

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

$$\pi_t = \beta \pi^e_{t+1|} + \kappa y_t^{gap} + u_t, \tag{A57}$$

which is equation (11).

D Analytical Results and Proofs

To see how lower attention weakens the effectiveness of forward guidance, consider the following stylized experiment.[8] The economy is hit by a negative natural rate shock in period $t = 0$ that pushes the nominal interest rate to the effective lower bound, i.e., $r_0^n < 0$ and $i_0 = -\underline{i}$. In $t = 1$, the natural rate returns to its steady state value and stays there indefinitely, $r_t^n = 0$ for all $t \geq 1$. From period $t = 2$ onwards, the output gap, and the real rate are back at their steady states, $y_t^{gap} = 0$ and $i_t - \pi_{t+1|t}^e = 0$ for all $t \geq 2$.

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

$$r_1 \equiv i_1 - \pi_{2|1}^e < 0 \tag{A58}$$

is the same for all γ and known in advance.[9] Hence, forward guidance here means to announce a certain value for the real rate. I discuss the implications of forward guidance via the nominal rate in section D.1. In the following, I assume that $(-\underline{i} - r_0^n + r_1)$ is negative, which means that the announced policy, captured by $r_1 < 0$, *makes up* for the binding lower bound in $t = 0$, captured by $-\underline{i} - r_0^n > 0$.

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

$$y_1^{gap} = -\varphi\left(r_1\right) > 0. \tag{A59}$$

Equation (A59) captures the *make-up policy:* by keeping the real rate below the natural rate, output is above potential after the lower-bound constraint stops to be binding.

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

$$y_0^{gap} = \underbrace{-\varphi\left(r_1\right)}_{=E_0 y_1^{gap}} - \varphi\left(-\underline{i} - \pi_{1|0}^e - r_0^n\right). \tag{A60}$$

Substituting the law of motion for inflation expectations

$$\pi^e_{1|0} = (1-\gamma)\pi^e_{0|-1} + \gamma\pi_0, \tag{A61}$$

into the Phillips Curve

$$\pi_0 = \frac{\beta}{1-\beta\gamma}(1-\gamma)\pi^e_{0|-1} + \frac{\kappa}{1-\beta\gamma}y_0^{gap} \tag{A62}$$

yields an expression for inflation expectations:

$$\pi^e_{1|0} = \frac{1-\gamma}{1-\beta\gamma}\pi^e_{0|-1} + \frac{\kappa\gamma}{1-\beta\gamma}y_0^{gap}. \tag{A63}$$

Putting everything together, we arrive at the following result.

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

$$y_0^{gap} = -\frac{\varphi\,(1-\beta\gamma)}{1-\gamma(\beta+\varphi\kappa)}\left[-\underline{i} - r_0^n + r_1\right] + \frac{\varphi(1-\gamma)}{1-\gamma(\beta+\varphi\kappa)}\pi^e_{0|-1} \tag{A64}$$

and inflation in $t = 0$ is given by

$$\pi_0 = -\frac{\kappa\varphi}{1-\gamma(\beta+\varphi\kappa)}\left[-\underline{i} - r_0^n + r_1\right] + (1-\gamma)\left[\frac{\beta}{1-\beta\gamma} + \frac{\varphi}{1-\gamma(\beta+\varphi\kappa)}\right]\pi^e_{0|-1}. \tag{A65}$$

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

Corollary 1. *Lower attention weakens*

(i) the negative effects of the shock,

(ii) the positive effects of forward guidance,

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

on the output gap and inflation.

Corollary 1 follows from the fact that the terms $\frac{\varphi(1-\beta\gamma)}{1-\gamma(\beta+\varphi\kappa)}$ and $\frac{\kappa\varphi}{1-\gamma(\beta+\varphi\kappa)}$ in front of $[-\underline{i} - r_0^n + r_1]$ are both increasing in γ. Points (i) and (ii) capture the main trade off of lower attention. While lower attention has a stabilizing effect via more anchored inflation expectations (point (i)), it renders forward guidance less effective (point (ii)). The reason why forward guidance becomes less effective as attention declines is because inflation expectations increase less in response to the announced policy, and thus, the real rate remains higher. Point (iii) illustrates an additional drawback of lower attention. A reduction of the effective lower bound, $-\underline{i}$, is less stimulating if agents in the economy are less attentive. Thus, going from a zero lower bound to a lower bound in negative territory, as conducted in several advanced economies over the last ten years, becomes less effective in terms of stimulating output and inflation if the public is inattentive (consistent with the exercise in figure 5). Away from the lower bound, point (iii) implies that the effectiveness of conventional monetary policy via the nominal interest rate becomes less effective as attention declines.

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

Corollary 2. *Lower attention*

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

$$(\beta + \varphi\kappa) > 1, \tag{A66}$$

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

$$\frac{\beta\,(\beta-1)}{(1-\beta\gamma)^2} + \frac{\varphi\,((\beta+\varphi\kappa)-1)}{(1-\gamma(\beta+\varphi\kappa))^2} > 0. \tag{A67}$$

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

Given the calibration in Table 4, conditions (A66) and (A67) both hold for all $\gamma < 0.99$. The effects of changes in γ, however, are numerically small. Thus, an increase in the average inflation rate—which increases average prior beliefs—is a promising monetary instrument to combat the loss of control via forward guidance as attention declines. By *ex-ante* increasing the average inflation rate, the policymaker not only supports higher inflation expectations and thus, lower real rates for a given nominal rate, but also gains additional policy space through the increase in the average nominal rate.

D.1 Extensions

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

$$\pi^e_{1|0} = (1-\rho_\pi)\bar{\pi} + \rho_\pi(1-\gamma)\pi^e_{0|-1} + \rho_\pi\gamma\pi_0, \tag{A68}$$

which can be substituted into the Phillips Curve:

$$\pi_0 = \frac{\beta}{1-\beta\rho_\pi\gamma}\left((1-\rho_\pi)\bar{\pi} + \rho_\pi(1-\gamma)\pi^e_{0|-1}\right) + \frac{\kappa}{1-\beta\rho_\pi\gamma}y_0^{gap}. \tag{A69}$$

Thus, inflation expectations are given by

$$\pi^e_{1|0} = \frac{1-\rho_\pi}{1-\beta\rho_\pi\gamma}\bar{\pi} + \frac{\rho_\pi(1-\gamma)}{1-\beta\rho_\pi\gamma}\pi^e_{0|-1} + \frac{\kappa\rho_\pi\gamma}{1-\beta\rho_\pi\gamma}y_0^{gap}. \tag{A70}$$

Putting everything together, we arrive at the following Proposition.

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

$$\begin{aligned} y_0^{gap} &= -\frac{\varphi\left(1-\beta\rho_\pi\gamma\right)}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}\left[-\underline{i}-r_0^n+r_1\right] \\ &+\frac{\varphi}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}\left[(1-\rho_\pi)\bar{\pi}+\rho_\pi(1-\gamma)\pi^e_{0|-1}\right] \end{aligned} \tag{A71}$$

and inflation in $t = 0$ is given by

$$\begin{aligned} \pi_0 = &-\frac{\kappa\varphi}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}\left[-\underline{i}-r_0^n+r_1\right]+(1-\rho_\pi)\left[\frac{\beta}{1-\beta\rho_\pi\gamma}+\frac{\varphi}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}\right]\bar{\pi} \\ &+\rho_\pi(1-\gamma)\left[\frac{\beta}{1-\beta\rho_\pi\gamma}+\frac{\varphi}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}\right]\pi^e_{0|-1}. \end{aligned} \tag{A72}$$

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

Corollary 3. *Lower attention*

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

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

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

Corollary 3 follows from the fact that the terms $\frac{\varphi(1-\beta\rho_\pi\gamma)}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}$ and $\frac{\kappa\varphi}{1-\rho_\pi\gamma(\beta+\varphi\kappa)}$ in front of $[-\underline{i}-r_0^n+r_1]$ are both increasing in γ. Points (*i*) and (*ii*) capture the main trade off of lower attention. While lower attention has a stabilizing effect via more anchored inflation expectations (point (*i*)), it renders forward guidance less effective (point (*ii*)). The reason why forward guidance becomes less effective as attention declines is because inflation expectations increase less in response to the announced policy, and thus, the real rate remains higher. Point (*iii*) illustrates an additional drawback of lower attention. A reduction of the effective lower bound, $-\underline{i}$, is less stimulating if agents in the economy are less attentive. Thus, going from a zero lower bound to a lower bound in negative territory, as conducted in several advanced economies over the last ten years, becomes less effective in terms of stimulating output and inflation if the public is inattentive. Note, that a decrease in the perceived inflation persistence, ρ_π, has the exact same implications as a decrease in γ.

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

Corollary 4. *Lower attention weakens the positive effects of higher long-run inflation beliefs* $\bar{\pi}$ *on output and inflation,*

Corollary 4 says that higher long-run beliefs have a positive effect on inflation and the output gap, but lower attention weakens these effects. However, as long as $\gamma\left(\beta+\varphi\kappa\right)<1$, a higher ρ_π mutes the effects of $\bar{\pi}$ on the output gap. Since this condition is usually satisfied and because ρ_π is in general close to 1, the role of high long-run inflation beliefs is quite weak. In the limit case $\rho_\pi \to 1$, long-run beliefs become irrelevant.

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

Corollary 5. *Lower attention*

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

$$\rho_\pi\left(\beta+\varphi\kappa\right)>1, \tag{A73}$$

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

$$\frac{\rho_\pi\beta\left(\rho_\pi\beta-1\right)}{\left(1-\beta\rho_\pi\gamma\right)^2}+\frac{\rho_\pi\varphi\left(\rho_\pi(\beta+\varphi\kappa)-1\right)}{\left(1-\rho_\pi\gamma(\beta+\varphi\kappa)\right)^2}>0. \tag{A74}$$

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

Given the calibration in Table 4, conditions (A73) and (A74) both hold. The effects of changes in γ, however, are numerically small. Thus, an increase in the average inflation rate—which also increases average prior beliefs—is a promising monetary instrument to combat the loss of control via forward guidance as attention declines. By *ex-ante* increasing the average inflation rate, the policymaker not only supports higher inflation expectations and thus, lower real rates for a given nominal rate, but also gains additional policy space through the increase in the average nominal rate. Higher average inflation, however, is also costly. In the analysis of optimal policy, later on, I will explore this trade off and characterize the optimal inflation target for different levels of attention.

D.1.1 Forward Guidance via Nominal Interest Rates

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

$$y_1^{gap} = -\varphi\left(i_1 - (1-\gamma)\,\gamma\pi_0 - \gamma\pi_1\right). \tag{A75}$$

The Phillips Curve in $t = 1$ yields

$$\pi_1 = \frac{(1-\gamma)\gamma}{1-\beta\gamma}\pi_0 + \frac{\kappa}{1-\beta\gamma}y_1^{gap}, \tag{A76}$$

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

$$y_1^{gap} = -\frac{\varphi\left(1-\beta\gamma\right)}{1-\gamma\left(\beta+\varphi\kappa\right)} i_1 + \varphi(1-\gamma)\gamma\frac{1+\gamma(1-\beta)}{1-\gamma(\beta+\varphi\kappa)}\pi_0. \tag{A77}$$

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

$$\pi_0 = \frac{\kappa}{1-\beta\gamma} y_0^{gap}, \tag{A78}$$

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

$$y_0^{gap} = \mathbb{E}_0 y_1^{gap} - \varphi\left(-\underline{i} - \frac{\kappa\gamma}{1-\beta\gamma} y_0^{gap} - r_0^n\right). \tag{A79}$$

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

LEMMA 6. *Forward guidance via the nominal interest rate yields the following output gap*

$$y_0^{gap} = A_1\left[-\frac{\varphi\left(1-\beta\gamma\right)}{1-\gamma\left(\beta+\varphi\kappa\right)} i_1 - \varphi\left(-\underline{i} - r_0^n\right)\right], \tag{A80}$$

and inflation

$$\pi_0 = \frac{\kappa}{1-\beta\gamma} A_1\left[-\frac{\varphi\left(1-\beta\gamma\right)}{1-\gamma\left(\beta+\varphi\kappa\right)} i_1 - \varphi\left(-\underline{i} - r_0^n\right)\right], \tag{A81}$$

where

$$A_1 \equiv \frac{1}{1-\varphi(1-\gamma)\gamma\frac{1+\gamma(1-\beta)}{1-\gamma(\beta+\varphi\kappa)}\frac{\kappa}{1-\beta\gamma} - \frac{\varphi\kappa\gamma}{1-\beta\gamma}}. \tag{A82}$$

Given the calibration in Table 4, A_1 is positive and increasing in γ. Thus, promising lower future nominal interest rates can indeed stimulate the economy. But similar to the case in which the policy maker commits to a certain future *real* rate, forward guidance becomes less effective when agents are less attentive. In fact, all three results from Corollary 3 go through.

Recall equation (A77):

$$y_1^{gap} = -\frac{\varphi\,(1-\beta\gamma)}{1-\gamma\,(\beta+\varphi\kappa)} i_1 + \varphi(1-\gamma)\gamma\frac{1+\gamma(1-\beta)}{1-\gamma(\beta+\varphi\kappa)}\pi_0. \tag{A83}$$

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

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

D.1.2 Heterogeneous Attention

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

LEMMA 7. *With heterogeneous attention to inflation, the output gap in $t = 0$ is given by*

$$y_0^{gap} = \frac{-\varphi\,(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H}\left[-\underline{i} + r_1 - r_0^n\right], \tag{A84}$$

and inflation by

$$\pi_0 = \frac{-\varphi\kappa}{1 - \beta\gamma_F - \kappa\varphi\gamma_H}\left[-\underline{i} + r_1 - r_0^n\right], \tag{A85}$$

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

Lemma 7 shows that a similar result as in Corollary 3 holds under heterogeneous attention levels.

Corollary 8. *Lower attention of either firms or households*

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

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

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

The parts concerning the output gap in Corollary 8 follow because the term in front of the brackets in equation (A84) becomes more negative as either of $\{\gamma_F, \gamma_H\}$ increases:

$$\frac{\partial\left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H}\right]}{\partial\gamma_F} = -\frac{\beta\kappa\varphi^2\gamma_H}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^2} < 0 \tag{A86}$$

$$\frac{\partial\left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H}\right]}{\partial\gamma_H} = -\frac{\varphi^2(1-\beta\gamma_F)}{(1-\beta\gamma_F-\kappa\varphi\gamma_H)^2} < 0, \tag{A87}$$

and the parts concerning inflation because the term $\frac{-\varphi\kappa}{1-\beta\gamma_F-\kappa\varphi\gamma_H}$ in equation (A85) becomes more negative as either of $\{\gamma_F, \gamma_H\}$ increases, too.

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

Corollary 9. *Lower levels of households' attention to inflation weaken the effectiveness of forward guidance, especially when firms' attention to inflation is low, and vice-versa.*

To see this, note that

$$\frac{\partial^2\left[-\frac{\varphi\kappa}{1-\beta\gamma_F-\kappa\varphi\gamma_H}\right]}{\partial\gamma_F\partial\gamma_H}=\frac{-2\varphi^2\kappa^2\beta}{\left(1-\beta\gamma_F-\kappa\varphi\gamma_H\right)^3}<0, \tag{A88}$$

$$\frac{\partial^2\left[\frac{-\varphi(1-\beta\gamma_F)}{1-\beta\gamma_F-\kappa\varphi\gamma_H}\right]}{\partial\gamma_F\partial\gamma_H}=\frac{-\beta\kappa\varphi^2\left[1-\beta\gamma_F+\kappa\varphi\gamma_H\right]}{\left(1-\beta\gamma_F-\kappa\varphi\gamma_H\right)^3}<0. \tag{A89}$$

D.2 Proof of Proposition 1

Proof. The New Keynesian Phillips Curve is given by

$$\pi_t = \beta \pi^e_{t+1|t} + \kappa y^{gap}_t + u_t. \tag{A90}$$

Substituting

$$\pi^e_{t+1|t} = \pi^e_{t|t-1} + \gamma \left(\pi_t - \pi^e_{t|t-1} \right) \tag{A91}$$

for $\pi^e_{t+1|t}$ yields

$$\pi_t = \beta \left(\pi^e_{t|t-1} + \gamma \left(\pi_t - \pi^e_{t|t-1} \right) \right) + \kappa y^{gap}_t + u_t \tag{A92}$$

$$\Leftrightarrow \pi_t (1 - \beta\gamma) = \beta \pi^e_{t|t-1} \left(1 - \gamma \right) + \kappa y^{gap}_t + u_t \tag{A93}$$

$$\Leftrightarrow \pi_t = \frac{\beta \pi^e_{t|t-1} \left(1 - \gamma \right) + \kappa y^{gap}_t + u_t}{\left(1 - \beta\gamma \right)} \tag{A94}$$

$$\Leftrightarrow \pi_t = \frac{\beta \left(1 - \gamma \right)}{\left(1 - \beta\gamma \right)} \pi^e_{t|t-1} + \frac{\kappa}{\left(1 - \beta\gamma \right)} y^{gap}_t + \frac{u_t}{\left(1 - \beta\gamma \right)}. \tag{A95}$$

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

E Additional Numerical Results

E.1 Non-Rational Output Gap Expectations

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

I then estimate attention to unemployment, γ^y, in the same way I estimate attention to inflation in Section 2, and I do so separately for the period before 1990 and the period after 1990. Table E14 shows the results. Attention to unemployment slightly increased from 0.088 before the 1990s ($\widehat{\gamma}_{y,<1990}$) to 0.100 after the 1990s ($\widehat{\gamma}_{y,\geq 1990}$). These differences, however, are not statistically significant, as the last column indicates. Similarly, when I set the break point at 2000, I estimate attention levels of 0.098 before 2000, and 0.099 after 2000. Again, the difference between the two is not statistically significantly different from 0. These results therefore indicate that while there was a strong decline in people's attention to inflation, their attention to unemployment did not change.

Now, to understand the policy implications of limited attention to the output gap, I

Table E14: Attention to unemployment

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

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

impose that output gap expectations are given by

$$y_{t+1|t}^{gap,e} = y_{t|t-1}^{gap,e} + \gamma^y \left(y_t^{gap} - y_{t|t-1}^{gap,e} \right) . \tag{A96}$$

With these expectations, the aggregate IS equation is given by

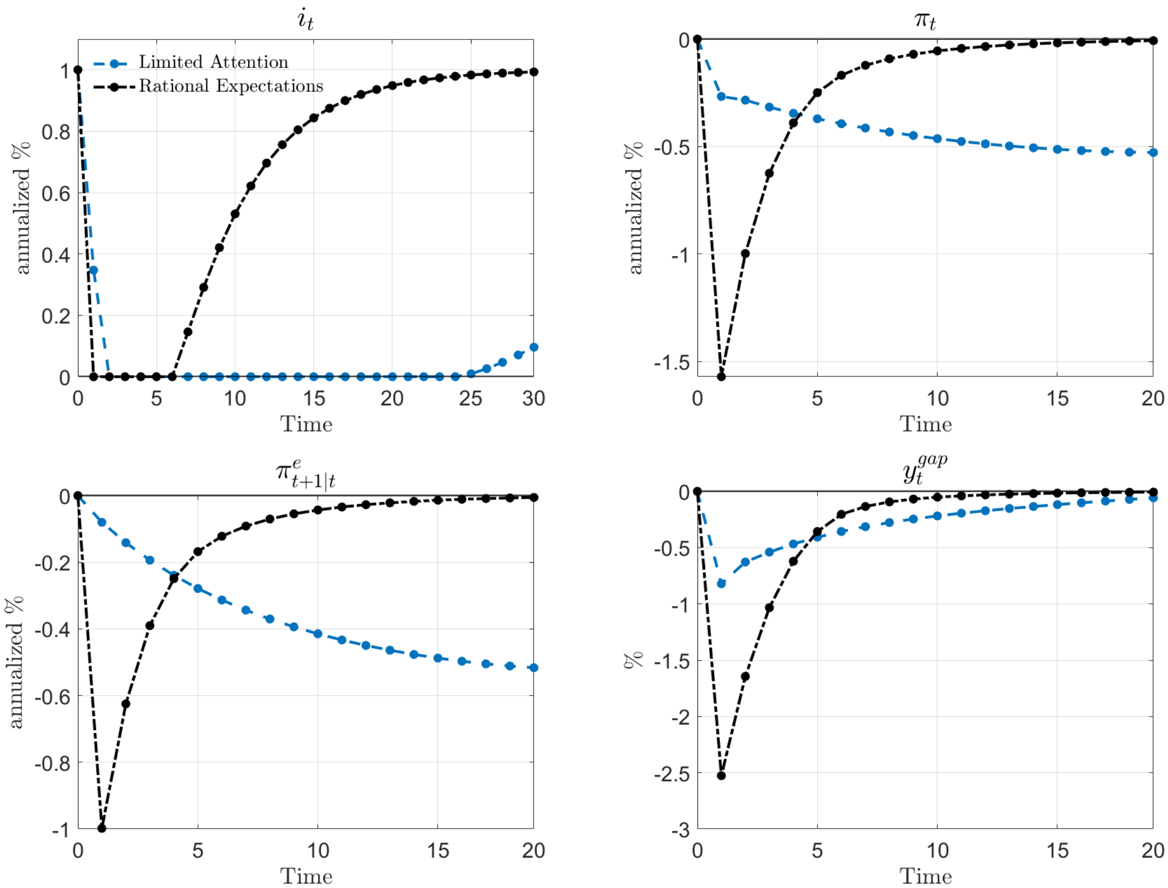
$$y_t^{gap} = y_{t+1|t}^{gap,e} - \varphi \left(i_t - \pi_{t+1|t}^e - r_t^n \right) , \tag{A97}$$

whereas the Phillips Curve and the Taylor rule remain unchanged.

Figure E5 shows the impulse response functions of the main variables in this economy after a negative three-standard deviation natural rate shock and with $\gamma^y = 0.1$ (I set γ^π to 0.3, as in Section 3 and keep the rest of the calibraton also unchanged). We see that the inflation-attention traps get exacerbated. The reason is that now make-up policies are even less effective because not only inflation expectations are backward looking but also output gap expectations. Thus, even though there is interest-rate smoothing in the Taylor rule which features some form of make-up policy, this is not effective in stimulating expectations and thus, the economy remains stuck at the ELB even longer. Furthermore, inflation, inflation expectations, and now also the output gap stay below their initial values very persistently.

Table E15 shows the implications of limited attention to the output gap for optimal policy. The upper part of the table shows the optimal inflation target and welfare when output gap expectations are rational, and the lower part shows the results when output gap expectations are non-rational. The table highlights the following two main results: (i)

Figure E5: Impulse Response Functions with Non-Rational Output Gap Expectations



Note: This figure shows the impulse-response functions of the nominal interest rate (upper-left panel), inflation (upper-right panel), inflation expectations (lower-left) and the output gap (lower-right) to a negative natural rate shock of three standard deviations in the case where output gap expectations are given by equation (A96). The blue-dashed lines show the case for the limited-attention model and the black-dashed-dotted lines for the rational expectations model. Everything is in terms of percentage deviations from the respective steady state levels, except the nominal rate is in levels.

limited attention to the output gap increases the optimal inflation target quite substantially and decreases welfare (independent of the level of γ^y or γ^π), and (ii) higher attention to the output gap reduces the optimal inflation target and increases welfare. The second result mirrors the main result regarding attention to inflation: lower attention (to inflation or the output gap) is welfare deteriorating in the presence of an effective lower bound constraint on nominal interest rates.

Table E15: Non-rational output gap expectations

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

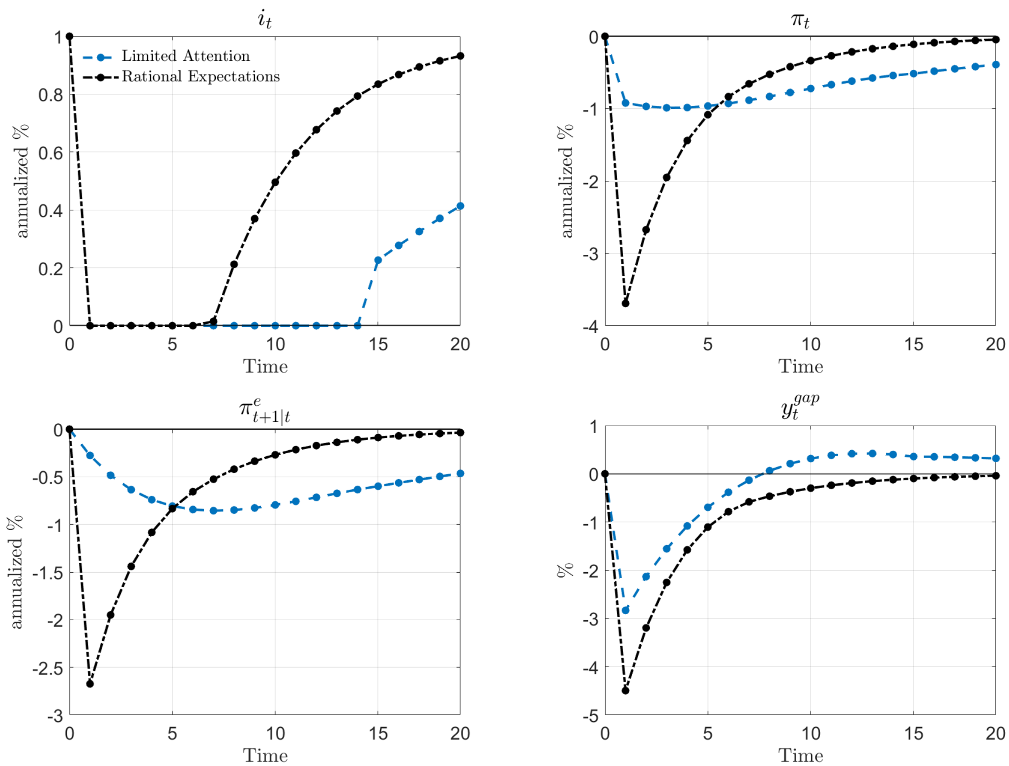
Notes: This table shows the implications of limited attention to the output gap for the optimal inflation target and welfare, for different combinations of γ^y (attention to the output gap) and γ^π (attention to inflation).

E.2 Different Taylor Rule

To show that the exact specification of the Taylor rule is not essential for the occurence of inflation-attention traps, Figure E6 shows the impulse-response functions of the nominal interest rate, inflation, inflation expectations and the output gap for the model in which the Taylor rule absent the ELB is given by

$$i_t = 1.5\pi_t. \tag{A98}$$

Figure E6: IRFs to Natural Rate Shock for Taylor rule (A98)

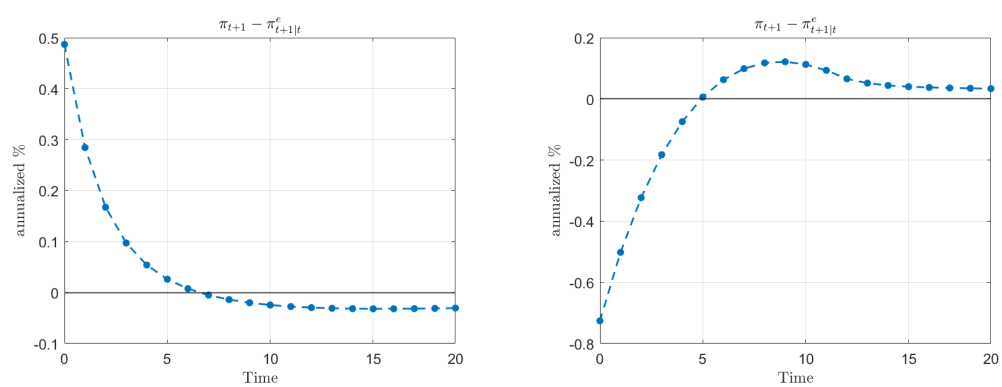


Note: This figure shows the impulse-response functions of the nominal interest rate (upper-left panel), inflation (upper-right panel), inflation expectations (lower-left) and the output gap (lower-right) to a negative natural rate shock of three standard deviations. The blue-dashed lines show the case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

E.3 Forecast Errors

Angeletos, Huo and Sastry (2021) propose a new test of models that deviate from FIRE. Namely, that expectations should initially underreact but overshoot eventually. A straightforward way to test this is to look at the model-implied impulse response functions of the forecast error, $\pi_{t+1} - \pi^e_{t+1|t}$, to an exogenous shock. Figure E7 shows these IRFs. The left panel shows the IRF of the forecast error after a positive natural rate shock and the right panel shows the corresponding IRF to a negative natural rate shock. In both cases, we see an underreaction in expectations, which manifests itself in a positive forecast error after a shock that increases the forecasted variable, and vice-versa following a negative shock. After about 5-6 periods, the forecast error response, however, flips sign. This is exactly the eventual overreaction, mentioned above and documented in Angeletos, Huo and Sastry (2021). Thus, my model of inflation expectations matches these empirical findings.

Figure E7: Impulse Response Functions of Forecast Errors



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

E.4 No Random Walk

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

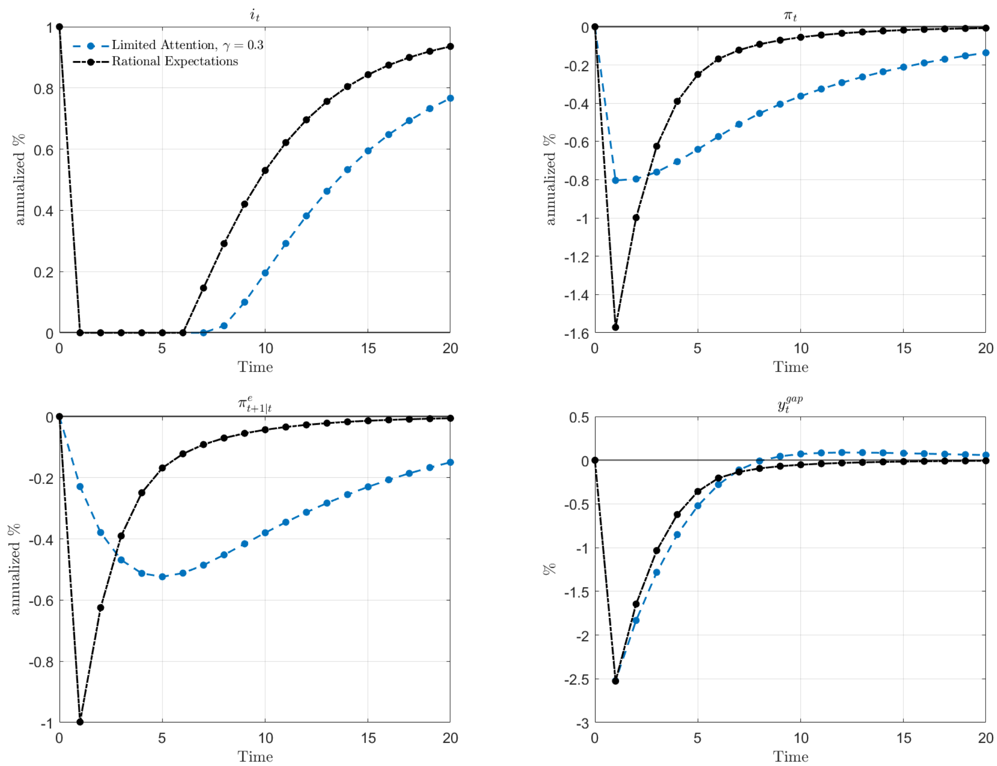
Figure E8 shows the same impulse response functions as reported in Figure 2 for the case of $\rho_\pi = 0.95$ and an average inflation of 0. We see a similar pattern, even though somewhat less pronounced. Inflation is persistently lower under limited attention due to slowly-adjusting inflation expectations. Expectations are updated even more sluggishly when $\rho_\pi < 1$. Further, this also dampens the initial response in inflation expectations, and thus, of inflation itself. Therefore, the attention trap is somewhat mitigated and the economy escapes the lower bound faster than with $\rho_\pi = 1$. Nevertheless, the nominal interest rate is low for longer due to the slow recovery of inflation.

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

$$\pi^e_{t+1|t} = (1-\rho_\pi)\bar{\pi} + \rho_\pi \pi^e_{t|t-1} + \rho_\pi \gamma \left(\pi_t - \pi^e_{t|t-1}\right), \tag{A99}$$

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

Figure E8: Impulse Response Functions to a Negative Natural Rate Shock

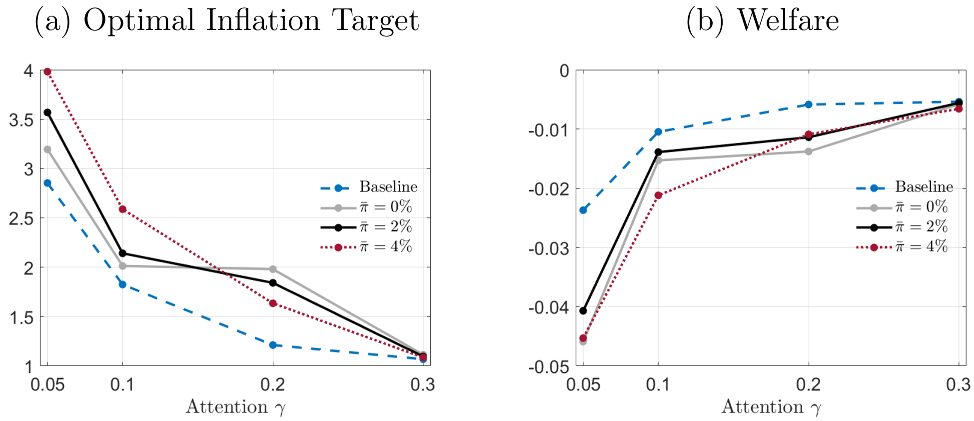


Note: This figure shows the impulse-response functions of the nominal interest rate (upper-right panel), inflation (upper-left panel), inflation expectations (lower-right) and the output gap (lower-left) to a negative natural rate shock of three standard deviations. The blue-dashed lines show the case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

Figure E9 shows the optimal inflation target (left panel) and welfare (17) (right panel) under Ramsey optimal policy for different levels of attention and different mean beliefs, $\bar{\pi}$. The blue-dashed lines show the results for the case with $\rho_\pi = 1$ (which is the baseline case discussed above), the gray-dashed-dotted lines show the results for $\rho_\pi = 0.95$ and $\bar{\pi} = 0\%$, the black-solid lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 2\%$, and the red-dotted lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 4\%$.

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

Figure E9: Mean Bias, Optimal Inflation Target and Welfare



Notes: This figure shows the average inflation rate under Ramsey optimal policy (left panel) and welfare (17) (right panel) under Ramsey optimal policy for different levels of attention and different mean beliefs, $\bar{\pi}$. The blue-dashed lines show the results for the case with $\rho_\pi = 1$ (which is the baseline case), the gray-dashed-dotted lines show the results for $\rho_\pi = 0.95$ and $\bar{\pi} = 0\%$, the black-solid lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 2\%$, and the red-dotted lines for $\rho_\pi = 0.95$ and $\bar{\pi} = 4\%$.

inflation expectations, actual inflation needs to increase more strongly. Therefore, a lower ρ_π can exacerbate attention traps when they occur.

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

$$\pi^e = \frac{(1 - \rho_\pi)\bar{\pi} + \rho_\pi \gamma \pi}{1 - \rho_\pi(1 - \gamma)}. \tag{A100}$$

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

LEMMA 10. *For the case $\rho_\pi = 1$, average inflation expectations move one-for-one with average inflation, independent of γ:*

$$\pi^e = \pi. \tag{A101}$$

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

$$0 < \frac{\partial \pi^e}{\partial \pi} = \frac{\rho_\pi \gamma}{1 - \rho_\pi (1 - \gamma)} < 1, \tag{A102}$$

and the strength of this dependency increases with γ

$$\frac{\partial^2 \pi^e}{\partial \pi \partial \gamma} > 0. \tag{A103}$$

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

$$0 < \frac{\partial \pi^e}{\partial \bar{\pi}} = \frac{(1 - \rho_\pi)}{1 - \rho_\pi (1 - \gamma)} < 1, \tag{A104}$$

and the strength of this dependency decreases with γ

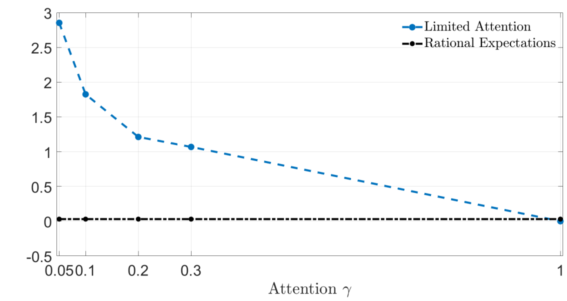
$$\frac{\partial^2 \pi^e}{\partial \bar{\pi} \partial \gamma} < 0. \tag{A105}$$

So, as attention falls, there are several opposing forces at work. On the one hand, the effect of $\bar{\pi}$ on average inflation expectations becomes stronger and thus, also exerts more pressure on actual inflation via the Phillips Curve. On the other hand, increasing the inflation target—average inflation—has a smaller effect on average inflation expectations at low levels of attention. Thus, to increase inflation expectations in this case, the inflation target needs to increase more strongly, which is of course costly. Comparing the optimal inflation targets in Figure E9, we see that at low levels of attention the first effect dominates. If $\bar{\pi}$ is relatively high, the inflation target is high.

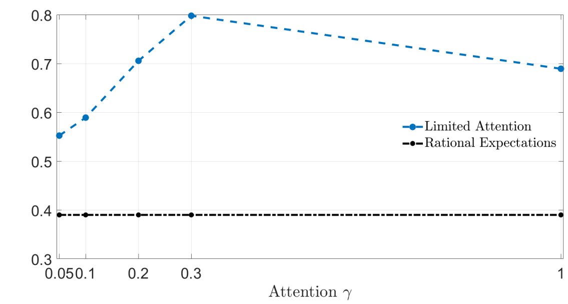
E.5 Figures to Section 4.3.5

Figure E10: Full Attention, $\gamma = 1$

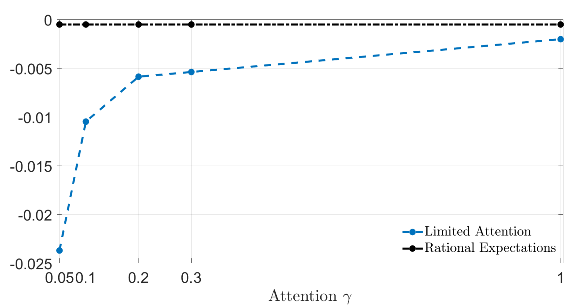
(a) Optimal Inflation Target



(b) Inflation Volatility



(c) Welfare



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

References

Adam, Klaus, and Mario Padula, "Inflation dynamics and subjective expectations in the United States," *Economic Inquiry*, 2011, *49* (1), 13–25.

Afrouzi, Hassan, Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar, "Overreaction in expectations: Evidence and theory," *The Quarterly Journal of Economics*, 2023.

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry, "Imperfect macroeconomic expectations: Evidence and theory," *NBER Macroeconomics Annual*, 2021, *35* (1), 1–86.

Ascari, Guido, and Lorenza Rossi, "Trend inflation and firms price-setting: Rotemberg versus Calvo," *The Economic Journal*, 2012, *122* (563), 1115–1141.

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, "Survey data and subjective beliefs in business cycle models," *Mimeo*, 2022.

Blundell, Richard, and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 1998, *87* (1), 115–143.

Hamilton, James D, *Time series analysis*, Princeton university press, 1994.

Matějka, Filip, and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 2015, *105* (1), 272–98.

Newey, Whitney K, and Kenneth D West, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation," *Econometrica*, 1987, *55* (3), 703–708.

Pfäuti, Oliver, and Fabian Seyrich, "A behavioral heterogeneous agent New Keynesian model," *Mimeo*, 2022.

Rotemberg, Julio J., "Sticky Prices in the United States," *Journal of Political Economy*, 1982, *90*, 1187–1211.

Weber, Michael, Stephen Sheflin, Tiziano Ropele, Rodrigo Lluberas, Serafin Frache, Brent Meyer, Saten Kumar, Yuriy Gorodnichenko, Dimitris Georgarakos, Olivier Coibion et al., "Tell Me Something I Don't Already Know: Learning in Low and High-Inflation Settings," *Mimeo*, 2023.