Inflation Persistence, Noisy Information and the Phillips Curve*

José-Elías Gallegos†

Banco de España

September 23, 2022

Abstract

A vast literature has documented that US inflation persistence has fallen in recent decades. However, this empirical finding is difficult to explain in monetary models. Using survey data on inflation expectations, I document a positive co-movement between ex-ante average forecast errors and forecast revisions (suggesting forecast sluggishness) from 1968 to 1984, but no co-movement afterwards. I extend the New Keynesian (NK) setting with noisy and dispersed information about the aggregate state, and show that inflation is more persistent in periods of greater forecast sluggishness. My results show that the change in firm forecasting behavior, documented in survey data, explains around 90% of the fall in inflation persistence since the mid 1980s. I also find that the changes in the dynamics of the Phillips curve in recent decades can be explained by the change in information frictions. Contrary to the literature which has emphasized a flattening of the NK Phillips curve in recent data, I do not find any evidence of the change in the slope of the Phillips curve once I control for the change in information frictions, but an increase in forward-lookingness. Finally, I find evidence of forecast underrevision in the post-COVID period, which explains the increase in the persistence of current inflation.

Keywords: Inflation persistence, Phillips curve, noisy information.

JEL Classifications: E31, E32, E52, E70.

*I am grateful to Tobias Broer, Alex Kohlhas, and Per Krusell for their advice and support. Further, I would like to thank Mattias Almgren, Gualtiero Azzalini, Timo Boppart, John Hassler, Ivan Jaccard, Markus Kondziella, John Kramer, Philip Lane, Kieran Larkin, Kurt Mitman, José Montalban, Athanasios Orphanides, Isabel Schnabel, Claire Thurwächter, Harald Uhlig, Joshua Weiss and seminar participants at Banco de España, Banca d'Italia, Bilkent, the CEBRA 2021 Annual Meeting, Copenhagen Business School, ECB Sintra Forum 2022, EEA-ESEM 2021 Congress, HEC Lausanne, Macro Finance Society, NOVA, RCEA Money, Macro & Finance Conference, Tilburg, UECE Conference on Economic and Financial Adjustments, Université Libre de Bruxelles, the IIES Macro Group and the IIES IMDb for useful feedback and comments. The views expressed are those of the author and do not necessarily represent the views of the Banco de España and the Eurosystem.

†Banco de España, Alcalá 48, Madrid, Spain. E-mail: jose.elias.gallegos@iies.bde.es

1 Introduction

Since long, expectations have played a central role in macroeconomics. However, most of work considers a limited theory of expectation formation, in which agents are perfectly and homogeneously aware of the state of nature and others' actions. In this paper, I consider a theory of expectation formation that incorporates significant heterogeneity and sluggishness in agents' forecasts, thus relaxing the standard *full information rational expectations* (FIRE) benchmark.[1] I include such expectation formation features into an otherwise standard New Keynesian (NK) model by introducing noisy and dispersed information, rationally processed separately by each agent, and match the information-specific parameters to the observed sluggishness in forecasts. I use this framework to interpret two empirical challenges in the literature: the fall in inflation persistence and the change in the dynamics of the Phillips curve.

As for the first empirical challenge, evidence suggests that the dynamic properties of US inflation have not been constant over time. In particular, inflation in the post-war period exhibits a high degree of persistence up until the mid 1980s, falling significantly since then. This fall in inflation persistence is not easily understood through the lens of monetary models, which has resulted in the "inflation persistence puzzle" (Fuhrer 2010).[2] This break coincides with a change in the US Federal Reserve's communication policy, which became more transparent and informative after the mid 1980s. Using survey data on US firms' forecasts, I document a significant sluggishness in responses to new information until the mid 1980s, but no evidence of sluggishness afterwards. The theoretical framework I build is consistent with this evidence. I argue that the change in the Fed communication improves firms' information and I use my model to show that the reduced stickiness in firms' inflation forecasts explains the fall in inflation persistence.

The second empirical challenge documents that the dynamics of the Phillips curve have changed in recent decades. The literature has mainly focused on the output gap coefficient, arguing for a flatter curve in recent decades (del Negro et al. 2020; Ascari and Fosso 2021). This finding indirectly implies that central bank actions, understood as nominal interest rate changes, are less effective in affecting inflation. I argue from the perspective of my model that the change in the dynamics of the Phillips curve can be explained by a lack of

1. I define sluggishness as the stickiness of current expectations on past expectations. I measure sluggishness as a positive co-movement between ex-ante average forecast errors and forecast revisions.

2. Persistence is an important property of a dynamic process since it determines both the memory of any past shock on today's outcome and its volatility. See Fuhrer (2010) for a handbook literature review.

backward-lookingness and an increase in forward-lookingness after the mid 1980s. I show that there is no evidence of a flattening in the Phillips curve once I control for the decline in information frictions.

I extend the textbook NK framework in Galí (2015) and Woodford (2003b) to noisy information following Lucas (1972), Woodford (2003a), Nimark (2008), Lorenzoni (2009), Huo and Takayama (2018), and Angeletos and Huo (2021). I assume that firms do not have complete and perfect information about the aggregate economic conditions. Instead, firms can observe their own granular conditions, the output they produce given their price, but they do not have perfect information about aggregate variables like inflation, output or interest rates. In place, they observe a noisy signal that provides information on the state of the economy, in this case the monetary policy shock. With this piece of information, firms form expectations on inflation, aggregate output and interest rates. This setting leads to a dynamic beauty contest in which firms need to form beliefs on what other firms believe about the economy. Morris and Shin (2002) and Woodford (2003a) are the first to study the economy as a static beauty contest, and Allen et al. (2006), Bacchetta and Van Wincoop (2006), Morris et al. (2006), and Nimark (2008) extend the economy to a dynamic beauty contest. More recently, Angeletos and Huo (2021) show that noisy information attenuates the general equilibrium effects associated with the Keynesian multiplier and the inflation-spending feedback, causing the economy to respond to news about the future *as if* agents were myopic. I extend the framework in Angeletos and Huo (2021) by merging the two blocks, the Dynamic IS and NK Philips curves, while still obtaining closed-form equilibrium dynamics that facilitate the interpretation of our results.[3]

In terms of the details of my model, I explain the fall in inflation persistence through a decrease in the degree of information frictions that firms face on central bank actions. Since the late 1960s, there has been a gradual improvement in the US Federal Reserve's public disclosure and transparency, sending clearer signals of their actions and future intentions to the market.[4] This has most notably occurred after 1985.[5] I show that in this framework,

3. Angeletos and Huo (2021) assume that firms observe the history of past price levels but do not extract any information from it, thus simplifying the framework. I assume that firms do not observe the price level.

4. See Lindsey (2003) for a comprehensive historical review.

5. Before 1967 the *Federal Open Market Committee* (FOMC), the US Fed decision unit, only announced policy decisions once a year in its Annual Report. In 1967, the FOMC decided to release the directive in the Policy Report (PR), 90 days after the decision. In 1976, the PR was enlarged and its delay was reduced to 45 days. Between 1976 and 1993 the information contained in the PR increased, without any further changes in the announcement delay. In 1977, the *Federal Reserve Reform Act* officially entitled the Fed with 3 objectives: maximum employment, stable prices and moderate long-term interest rates. In 1979, the first macroeconomic forecasts on real GNP and GNP inflation from FOMC members were made available. The

inflation is more persistent in periods of greater forecast sluggishness. Noisy information generates an underreaction to new information because individuals shrink their forecasts towards prior beliefs when the signals they observe are noisy. This endogenous anchoring in forecasts causes firms to set prices to their existing prior, thus slowing the speed of price changes. Using micro-data on inflation expectations from the Survey of Professional Forecasters (SPF) and the Livingston Survey on Firms, I document that firms' forecasts used to react sluggishly before the mid 1980s. However, there appears to be a break, and there is no evidence of sluggishness in recent decades. My results suggest that agents became more informed about inflation after the change in the Federal Reserve disclosure policy in the mid 1980s. Because inflation depends on the expectations of future inflation, the change in expectation formation feeds into inflation dynamics, which endogenously reduces inflation persistence. I find that this change in firms' forecasting behavior explains around 90% of the fall in inflation persistence since the mid 1980s.

I also study the dynamics of the Phillips curve over time through the lens of my model. The literature has mainly focused on the output gap coefficient, arguing for a flatter curve in recent decades, implying a fall in the sensitivity of inflation and the real side of the economy ("inflation disconnect" puzzle, see e.g., del Negro et al. 2020; Ascari and Fosso 2021). In the standard model, inflation dynamics are reduced to the NK Phillips curve, which relates current inflation to the current output gap and expected future inflation. Inflation is only related to the real side of the economy through the Phillips curve slope, and the only possible explanation for the lack of dependence of inflation on output in recent decades is a fall in the slope.

The literature has extensively focused on this slope, in the hope of documenting that this relation has weakened and that the inflation process is therefore largely independent of any change from the demand side of the economy, including changes in the policy rate or central bank actions. Armed with the noisy information framework, I find that the changes in the dynamics of the Phillips curve can be explained by the change in information frictions. First, I show that the NK Phillips curve is enlarged with a backward-looking

"tilt" (the likelihood regarding possible future action) was introduced in the PR in 1983. Between 1985 and 1991, the Fed introduced the "ranking of policy factors", which after each meeting ranked aggregate macro variables in importance, signaling priorities with regard to possible future adjustments. The minutes, a revised transcript of the discussions during the meeting, started being released together with the PR in 1993, 45 days after the meeting. In 1994 the FOMC introduced the immediate release of the PR after a meeting if there had been a change, coupled with an immediate release of the "tilt" (likelihood regarding possible future action) since 1999. Since January 2000 there has been an immediate announcement and press conference after each meeting, regardless of the decision.

term on lagged inflation and myopia towards expected future inflation. Once I correct for the misspecification in the NK Phillips curve, there is no evidence of a fall in its slope, but evidence of a reshuffle from a backward-looking curve to a more forward-looking curve. Second, I show that under a *general* information structure, the Phillips curve is modified such that current inflation is related to current and future output through two different channels: the output gap coefficient *and* firms' expectation formation process. I show that there is no empirical evidence of any change in the slope once I control for a decline in information frictions, using SPF forecasts. In summary, contrary to the literature which has emphasised a flattening of the NK Phillips curve in recent data, I do not find any evidence of the change in the structural slope once I control for imperfect expectations.

Roadmap The paper proceeds as follows. Section 2 documents the two empirical challenges and the decrease in forecast sluggishness and information frictions in recent decades. In Section 3, I describe the theoretical framework, and I derive the main results in section 4. Section 5 concludes the paper.

In Appendix D I revisit different theories that produce a structural relation between inflation and other forces in the economy, and I show that they cannot explain the fall in inflation persistence. In the benchmark NK model, inflation inherits its properties of the exogenous driving forces. Hence, in order to explain the fall in inflation persistence documented in the data, a fall in the persistence of these exogenous shocks is required. I find that the persistence of exogenous monetary policy, total factor productivity and other shocks has been remarkably stable in the post-war period. Acknowledging the fact that purely forward-looking models cannot generate intrinsic persistences, I extend the benchmark and explore backward-looking frameworks. I find that they generate little endogenous persistence, insufficient to generate the significant fall in inflation persistence that I observe in the data.[6]

2 Empirical Challenges and Information Frictions

In this section, I discuss the two empirical challenges and the change in inflation forecast underrevision. First, I provide empirical evidence on the fall in the persistence of inflation in recent decades. Second, I show that the change in the persistence coincides with a decline in forecast underreaction in the recent decades. Third, I argue that the documented changes

6. I extend the setting to price indexation, trend inflation and optimal monetary policy under discretion and commitment. I show that these frameworks cannot explain the large fall in inflation persistence.

	1968:Q4–2020:Q1	1968:Q4–1984:Q4	1985:Q1–2020:Q1
Mean	3.362	6.160	2.117
Volatility	2.400	2.234	1.016
First-Order Autocorrelation	0.880	0.754	0.505

Table I: Summary statistics over time.

in the persistence of inflation and forecast underreaction can explain the changes in the dynamics of the Phillips curve over time.

2.1 The First Puzzle: Inflation Persistence

A vast literature has documented that US inflation persistence has fallen in recent decades. Fuhrer and Moore (1995), Cogley and Sbordone (2008), Fuhrer (2010), Cogley et al. (2010), and Goldstein and Gorodnichenko (2019) find evidence of a structural break in the first-order autocorrelation of inflation in the 1980-1985 window, with persistence falling from around 0.75-0.8 to 0.5.[7] In this section I revisit this empirical challenge and document a fall in inflation persistence since the mid 1980s.[8] I use the (annualized) quarterly growth in the GDP Deflator as a proxy for aggregate inflation.[9]

The inflation time series is reported in Figure 1. I follow Fuhrer (2010) and divide the sample into two sub-periods, pre- and post-1985:Q1 until 2020:Q1. I report the mean and 2 standard deviation bands by each subperiod. Inflation started its upward trend in the 1960s, continuing in the next decade with two local peaks in the mid 1970s and in the early 1980s. Then, inflation started its downward trend lasting until the early 1990s, and has roughly remained at 2% afterwards. Differentiating between the two subperiods, one can see from the previous figure that the level of inflation has fallen from 6% to 2%, and that inflation has become less volatile.[10]

7. In a cross-country analysis, Benati and Surico (2008) find that countries with central banks that follow an inflation targeting policy experience lower persistence.

8. Inflation data is available at a quarterly frequency since 1947:Q1. However, I will stick to the 1968:Q4-2020:Q2 sample since I seek to link the results presented in this section to surveys on expectations, which are available since 1968:Q4.

9. I define the inflation rate at time t, π_t, as the (annualized) log growth in the index, $400 \times (\log X_t - \log X_{t-1})$, where X_t is the GDP deflator at time t.

10. I omit the fall in average inflation and volatility from the analysis, since both can be easily explained in a trend-inflation NK setup through a decrease in the inflation target of the central bank, and an increase in the aggressiveness of the monetary authority towards the inflation gap, for which Clarida et al. (2000) provide empirical evidence.

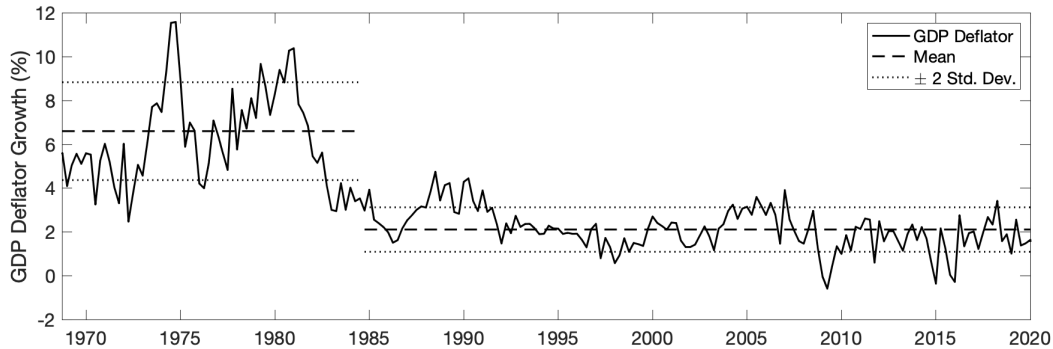


Figure 1: Time series of inflation, with subsample mean and standard deviation.

In the monetary literature, inflation is generally assumed to follow an independent autoregressive stochastic process. In such a case, the stationary mean depends both on the intercept and the lagged inflation coefficients. On the other hand, the stationary volatility depends both on the innovation volatility and the lagged inflation coefficients. Table I reports summary statistics on the mean, volatility and first-order autocorrelation by each subsample. In the following, I seek to investigate if these differences across subsamples are statistically significant.

Let us assume that inflation follows a simple AR(1) process with a drift. Recall that the change in the average level of inflation, documented in Figure 1, can be explained by two parameters, the intercept and the persistence coefficient. Once more, I follow Fuhrer (2010) and assume that the break date is 1985:Q1.[11] I test for a (potential) structural break on the intercept and persistence coefficients by estimating (2.1). Formally, I consider the regression

$$\pi_t = \alpha_\pi + \alpha_{\pi,*}\mathbb{1}_{\{t\geq t^*\}} + \rho_\pi \pi_{t-1} + \rho_{\pi,*}\mathbb{1}_{\{t\geq t^*\}}\pi_{t-1} + e_t \tag{2.1}$$

where $\mathbb{1}_{\{t\geq t^*\}}$ is an indicator variable equal to 1 if the period is within the post-1985 era, and e_t is the error term. The advantage of relying on a specification like (2.1) instead of a cross-sample analysis as in Table I is that the former allows us to verify if the structural change in the coefficients is statistically significant. I report my findings in Table II. First, I find that both the intercept and the persistence are highly significant when I consider the full sample with no structural break (column 1). Second, I find strong evidence of a structural

11. To confirm the break date, I additionally test for the null of no structural break in inflation dynamics around 1985:Q1. I reject the null of no break (p-value = 0.000). If I instead am agnostic about the break date(s), the test suggests that the break occurred in 1991:Q1, with the lower and upper 95% confidence bands 1986:Q1 and 1996:Q1.

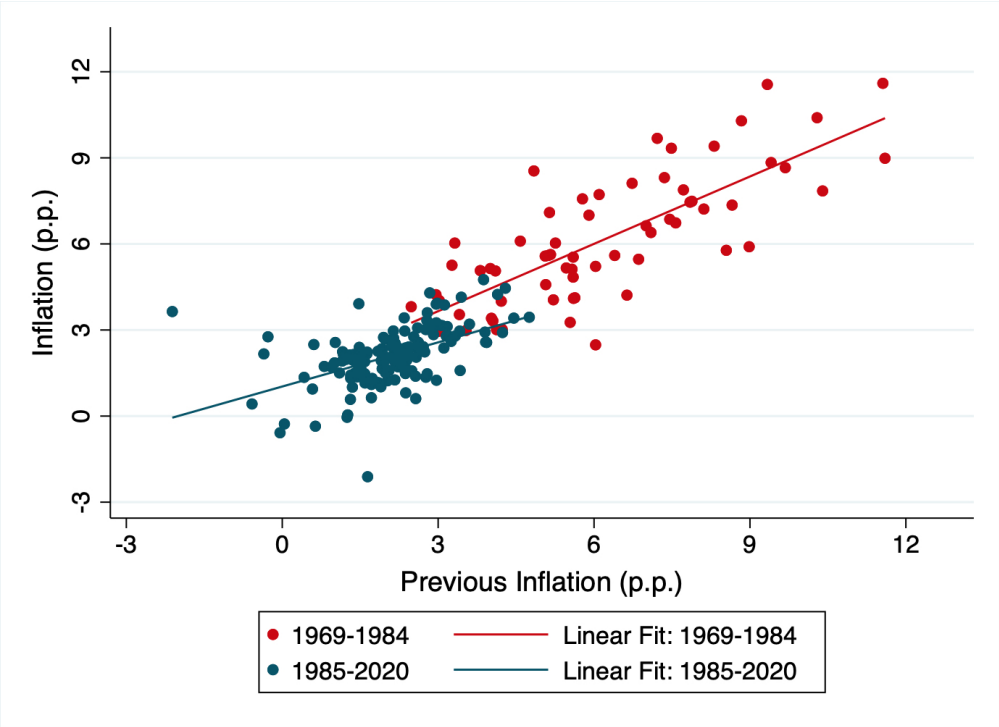


Figure 2: Scatter plot contemporaneous inflation (vertical axis) and one-quarter lagged inflation (horizontal axis). Red dots correspond to 1968-1984 observations, and blue dots correspond to observations after 1984.

	(1) Full Sample	(2) Structural Break
π_{t-1}	0.880***	0.785***
	(0.0466)	(0.0755)
$\pi_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$		-0.287**
		(0.144)
Constant	0.400**	1.320***
	(0.166)	(0.471)
Constant$\times \mathbb{1}_{\{t \geq t^*\}}$		-0.263
		(0.543)
Observations	206	206

HAC robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table II: Estimates of regression (2.1)

break in persistence, falling from 0.79 in the pre-1985 period to 0.5 afterwards (column 2). On the other hand, I do not find any evidence of a structural break in the intercept. This structural break is also easily visualized in the scatter plot in Figure 2. Considering these findings, and the robustness checks discussed in Appendix B.1, I conclude that (*i*) inflation persistence has fallen since 1985, and that (*ii*) the fall in the level of inflation documented in Figure 1 is explained through a change in the persistence of inflation.[12]

2.2 Evidence on Information Frictions

As discussed in the introduction, the actions of the Fed have become more transparent over time. The delay between the Fed's action and the announcement to the public has been shortened from around a year to a few minutes, and the amount of information contained in the Policy Report and other documents released to the public has increased substantially.[13] In this section, I document a contemporaneous change in beliefs and expectation formation around the same date in which inflation persistence is reported to break. Using survey data on US firms' forecasts, I document a significant sluggishness in responses to new information until the mid 1980s, but no evidence of sluggishness afterwards. Using expectations data from the Survey of Professional Forecasters (SPF), I study whether there is a significant change in different measures of information frictions around 1985:Q1.[14]

The problem that the econometrician faces when trying to quantify or estimate the degree of information frictions is that she does not know what each agent, or the average agent, has observed at any given point in time. The literature has approached this regression design problem by measuring the change in actions after an inflow of information. Consider, for example, the average forecast of annual inflation at time t, $\overline{\mathbb{E}}_t \pi_{t+3,t}$, where $\pi_{t+3,t}$ is the GDP deflator growth between periods $t+3$ and $t-1$. One can think of this object as the action that the average forecaster makes. Let us now consider the average forecast of 4-quarters-ahead inflation at time $t-1$, $\overline{\mathbb{E}}_{t-1} \pi_{t+3,t}$. The difference between these two objects, the average forecast revision, $\text{revision}_t \equiv \overline{\mathbb{E}}_t \pi_{t+3,t} - \overline{\mathbb{E}}_{t-1} \pi_{t+3,t}$, provides us with information about the

12. I explore alternative analyses, obtaining similar results, in Appendix B.1. I consider (*i*) two alternative measures of inflation, price inflation (CPI) and producer inflation (PCE), (*ii*) rolling-sample and time-varying estimates, and (*iii*) a subsample unit root analysis.

13. I provide a more detailed historical analysis of the Fed's gradual increase in transparency in Appendix E.

14. The American Statistical Association and the National Bureau of Economic Research started the survey in 1968:Q4, which has been conducted by the Federal Reserve Bank of Philadelphia since 1990:Q1. Every three months, professional forecasters are surveyed on their forecasts on economic variables like output, inflation or interest rates. These forecasters work at Wall Street financial firms, commercial banks, consulting firms, university research centers and other private sector companies.

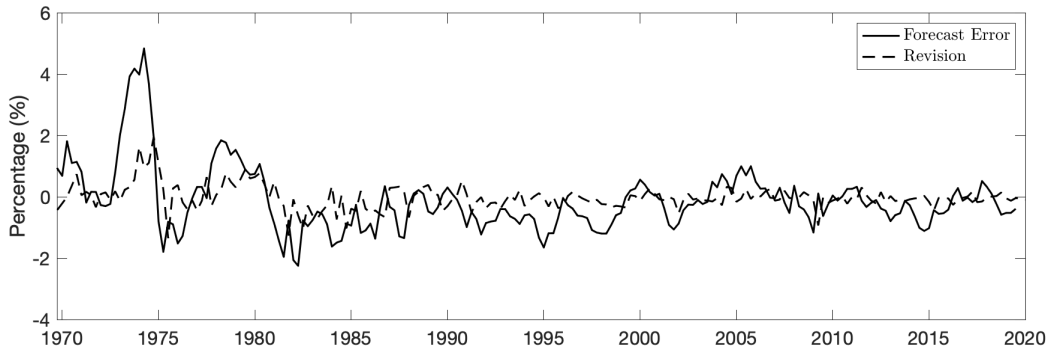


Figure 3: Time series of ex-ante average forecast errors and forecast revisions.

average agent action after the inflow of information between periods t and $t-1$. I plot the raw series in Figure 3. Recent research (Coibion and Gorodnichenko 2012, 2015a) has documented a positive co-movement between ex-ante average forecast errors, denoted by forecast error$_t \equiv \pi_{t+3,t} - \overline{\mathbb{E}}_t \pi_{t+3,t}$, and average forecast revisions.[15] Formally, their regression design is

$$\text{forecast error}_t = \alpha_{\text{rev}} + \beta_{\text{rev}} \text{ revision}_t + u_t, \tag{2.2}$$

where a positive co-movement ($\hat{\beta}_{\text{rev}} > 0$) suggests that positive revisions predict positive forecast errors.[16] That is, after a positive revision of annual inflation forecasts, agents consistently under-predict inflation. Although I only focus on firms in the main text, this form of forecast stickiness or sluggishness is consistent across different agent types (see Coibion and Gorodnichenko 2012, 2015a for evidence on consumers, firms, central bankers, etc.)

15. I use the first-release value of annual inflation, since forecasters do not have access to future revisions of the data when they provide their forecast.

16. Under the FIRE assumption, β_{rev} should be zero. Each agent's individual forecast is identical to each other agent's forecast. As a result, the average expectation operator in (2.2) could be interpreted as a representative agent forecast, and one would be effectively regressing the forecast error of the representative agent on its forecast revision. Under RE, the forecast revision should not consistently predict the forecast error. Otherwise, the agent would incorporate this information in his information set. Therefore, a positive estimate of β_{rev} in the above regression suggests that the FIRE assumption is violated. In this model, I maintain the RE assumption, and assume that agents face information frictions, thus generating heterogenous beliefs (information sets) across households. Bordalo et al. (2018) and Broer and Kohlhas (2019) find evidence of a violation of the rational expectations assumption by regressing (2.2) at the individual level, finding evidence of agent over-confidence when forecasting inflation. Notice that even if I assume information frictions, the above regression at the individual level should report a β_{rev} estimate of zero, because at the individual level the forecast revision should not consistently predict the forecast error. I do not assume a departure from rational expectations because, as shown in Angeletos et al. (2020), over-confidence would have no effect on aggregate dynamics and would therefore not affect the inflation persistence.

The results, reported in the first column in Table III, suggest a strong violation of the FIRE assumption: the measure of information frictions, β_{rev}, is significantly different from zero. Agents underrevise their forecasts: a positive β_{rev} coefficient suggests that positive revisions predict positive (and larger) forecast errors. In particular, a 1 percentage point revision predicts a 1.23 percentage point forecast error. The average forecast is thus smaller than the realized outcome, which suggests that the forecast revision was too small, or that forecasts react sluggishly.

Following the previous analyses on inflation persistence, I assume that the break date is 1985:Q1.[17] Following a similar structural break analysis as in Section 2.1, I study if there is a change in expectation formation around the same break date. Formally, I test for a structural break in belief formation around 1985:Q1 by estimating the following structural-break version of (2.2),

$$\text{forecast error}_t = \alpha_{\text{rev}} + \beta_{\text{rev}}\text{revision}_t + \beta_{\text{rev}*}\mathbb{1}_{\{t \geq t^*\}}\text{revision}_t + u_t \tag{2.3}$$

A significant estimate of $\beta_{\text{rev}*}$ suggests a break in the information frictions. The results in the fourth and fifth columns in Table III suggest that there is a structural break around 1985:Q1. The estimate $\hat{\beta}_{\text{rev}*} < 0$ suggests that firms' forecasts underreact less since 1985 (in fact, I do not find any evidence of forecast stickiness.) This structural break finding is also easily visualized in the scatter plot in Figure 4.

In the lens of a noisy and dispersed information framework, this implies that agents became *more* more informed about inflation, with individual forecasts relying less on priors and more on news.[18] These structural break findings are consistent with alternative measures of information frictions, as discussed in Appendix B.4.[19]

17. I test for the null of no structural break in inflation dynamics around 1985:Q1. I reject the null of no break (p-value = 0.01). If I am instead agnostic about the break date(s), the test suggests that the break occurred in 1980:Q1.

18. The forecast underreaction behavior is consistent with many different FIRE extensions of the benchmark setting. In Appendix B.4, I show that alternative moments in survey data are only consistent with noisy and dispersed information. I find that ex-ante inflation forecasts errors react to monetary shocks before 1985 and not after, and that the cross-sectional volatility of inflation forecasts does not react to monetary shocks throughout the sample.

19. I conduct robustness checks studying the impulse response of ex-ante inflation forecast errors to ex-ante monetary policy shocks, the cross-sectional volatility of inflation forecasts over time, or using alternative datasets like the Livingston Survey.

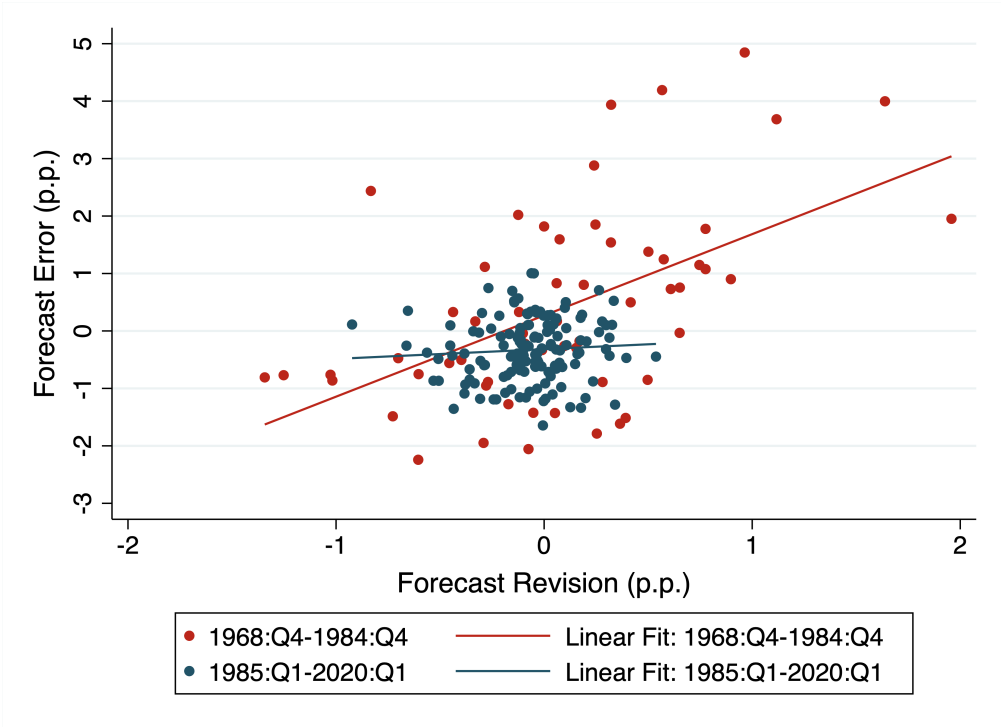


Figure 4: Scatter plot of ex-ante average forecast error (vertical axis) and average forecast revisions (horizontal axis). Red dots correspond to 1968-1984 observations, and blue dots correspond to observations after 1984.

	Full Sample	1968:Q4-1984:Q4	1985:Q1-2020:Q1	Structural Break	
	(1)	(2)	(3)	(4)	(5)
Revision	1.230***	1.414***	0.169	1.501***	1.414***
	(0.250)	(0.283)	(0.193)	(0.317)	(0.281)
Revision $\times$ $\mathbb{1}_{\{t \geq t^*\}}$				-1.111***	-1.245***
				(0.379)	(0.341)
Constant	-0.0875	0.271	-0.317***	-0.135*	0.271
	(0.0696)	(0.185)	(0.0478)	(0.0690)	(0.184)
Constant $\times$ $\mathbb{1}_{\{t \geq t^*\}}$					-0.587***
					(0.190)
Observations	197	58	139	197	197

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table III: Estimates of regression (2.3)

2.3 The Second Puzzle: The Phillips Curve

Unemployment has fluctuated between historically large and low levels since 1985. During the Great Recession, unemployment increased to a level comparable to that of the Volcker disinflation. Shortly after that, unemployment decreased to unprecedented low levels. Throughout this period, inflation seemed to be unaffected and disconnected from the changes in the real side of the economy, with no disinflation during the Great Recession and no large inflation afterwards (Hall 2011; Ball and Mazumder 2011; Coibion and Gorodnichenko 2015b; del Negro et al. 2012; Lindé and Trabandt 2019). This contrasts with the Volcker disinflation experience, which caused a large increase in unemployment and gave rise to the concept of the sacrifice ratio.[20]

Taking a model-oriented view, this second empirical challenge implies that the Phillips curve has flattened in recent decades, implying that inflation is no longer affected by other real variables (del Negro et al. 2020; Ascari and Fosso 2021; Atkeson and Ohanian 2001; Stock and Watson 2019).[21] The most well-known (structural) inflation equation is the NK Phillips curve,

$$\pi_t = \kappa \tilde{y}_t + \beta \mathbb{E}_t \pi_{t+1} \tag{2.4}$$

which relates current inflation π_t to the current output gap $\tilde{y}_t$ and expected future inflation $\mathbb{E}_t \pi_{t+1}$. Notice that, in this framework, inflation is *only* related to output *through* the Phillips curve slope κ. In such framework, the most prominent explanation for the lack of dependence of inflation on output is the fall in κ.[22] The literature has extensively focused on this coefficient, in the hope of showing that this relation has somehow flattened and that inflation is less dependent on any other variable. The available empirical evidence is mixed, with most recent evidence arguing for a (small and) constant slope over time (McLeay and Tenreyro 2020; Hazell et al. 2020).

In section 4.2, I argue that an extension to the benchmark model, in which the assumption of complete and full information is relaxed, enlarges the Phillips curve (2.4) with anchoring and myopia. I argue that the finding that the slope of the Phillips curve has fallen in recent decades is simply the result of a misspecified Phillips curve equation (2.4). In particular, I

20. The sacrifice ratio measures the change in output per each 1% change in inflation.

21. This finding indirectly implies that central bank actions, understood as nominal interest rate changes, are less effective in affecting inflation.

22. Another explanation put forward by McLeay and Tenreyro (2020) is that a monetary authority conducting optimal monetary policy under discretion could explain the disconnect without resorting to κ. Although appealing, I find that this change cannot explain the contemporaneous fall in persistence.

argue from the perspective of my model that the change in the dynamics of the Phillips curve can be explained by a lack of backward-lookingness and an increase in forward-lookingness after the mid 1980s, which is supported by the data. Once these additional terms have been controlled for, and I estimate a Phillips curve closer to the hybrid version implied by price-indexation settings, I do not find any evidence of a change in κ.

3 Noisy Information

In this section, I consider a theory of expectation formation that incorporates significant heterogeneity and sluggishness in agents' forecasts, thus relaxing the standard *full information rational expectations* (FIRE) benchmark. I include such expectation formation features into an otherwise standard New Keynesian (NK) model by introducing noisy and dispersed information, rationally processed separately by each agent, and match the information-specific parameters to the observed sluggishness in forecasts. I will argue that the change in the Fed communication improved firms' information, and I use my model to show that the reduced underreaction in firms' inflation forecasts will translate into reduced persistence in inflation.[23,24] I show that in this framework, inflation is more persistent in periods of greater forecast sluggishness. Noisy information generates an underreaction to new information because individuals shrink their forecasts towards prior beliefs when the signals they observe are noisy. This endogenous anchoring in forecasts causes firms to set prices to their existing prior, thus slowing the speed of price changes. Because inflation depends on the expectations of future inflation, the change in expectation formation feeds into inflation dynamics, which endogenously reduces inflation persistence. I find that this change in firm forecasting behavior explains around 90% of the fall in inflation persistence since the mid 1980s.

I discuss a variety of New Keynesian models in Appendix D, and show that none of them can produce a significant fall in inflation persistence. The intuition behind that result is that, in purely forward frameworks, inflation is proportional to the exogenous shocks,

23. The delay between the Federal Reserve action and the announcement to the public has been shortened from around a year to a few minutes and there has been a substantial increase in the amount of information contained in the PR and other documents released to the public has substantially increased. I provide a more detailed historical analysis of the Fed's gradual increase in transparency in Appendix E.

24. A criticism to the gradual information disclosure argument is that, although actions *themselves* could not be known with any certainty until after a year, market participants could observe the changes in interest rates and monetary aggregates induced by the action and could thus infer the action, in the spirit of the Grossman and Stiglitz (1980) paradox. To alleviate this concern, I measure information frictions using data from professional forecasters. The underlying assumption here is that professional forecasters are among the most informed agents in the economy since their job is to make predictions for private companies. Obtaining evidence on significant information frictions would therefore invalidate the previous criticism.

and only extrinsically persistent. I show that the persistence of these exogenous shocks has not changed over time. Then, I explore several extensions that produce backward-looking dynamics, such as optimal monetary policy under commitment, price indexation or positive trend inflation. I argue that these extensions generate mild anchoring and cannot explain the documented *change* in inflation persistence.

3.1 Noisy Information New Keynesian Model

In order to relate the previous empirical findings on inflation persistence to information frictions, I build a noisy information New Keynesian model based on the island setting by Lucas (1972), Woodford (2001), Nimark (2008), Lorenzoni (2009), and Angeletos and Huo (2021).[25] Firms observe the economic conditions in their island, but they do not have full information about the economic conditions in the archipielago. In particular, firms can observe their own granular conditions, such as their production given their price, but they do not have perfect information about aggregate macro variables like inflation, output or interest rates. They observe a noisy signal that provides information on the state of the economy, in this case the monetary policy shock. With this piece of information, firms form expectations on inflation, aggregate output and interest rates. For simplicity, I assume that households and the monetary authority have access to full information.[26]

Apart from this information friction, which I describe formally below, firms are subject to the standard Calvo-lottery price friction, which allows us to write the price-setting problem as a forward-looking one, and compete in a monopolistic economy. There is a continuum of firms indexed by $j \in \mathcal{I}_f = [0, 1]$, each being a monopolist producing a differentiated intermediate-good variety with CES ϵ, producing output Y_{jt} and setting price P_{jt}. Technology is represented by the production function

$$Y_{jt} = N_{jt}^{1-\alpha} \tag{3.1}$$

where $1 - \alpha$ is the labor share.

Aggregate Price Dynamics As in the benchmark NK model, price rigidities take the form of a Calvo-lottery. In every period, each firm can reset its price with probability $(1-\theta)$, independent of the time of the last price change. That is, only a measure $(1 - \theta)$ of firms

25. The derivation of the model is relegated to Online Appendix F.
26. I relax the FIRE assumption on households in Appendix C.

is able to reset their prices in a given period, and the average duration of a price is given by $1/(1-\theta)$. Let $p_t = \log P_t$ denote the (log) aggregate price level and $p_t^* = \log P_t^*$ the (log) aggregate price set by firms which are able to act. Such an environment implies that aggregate price dynamics are given (in log-linear terms) by

$$p_t = (1-\theta)p_t^* + \theta p_{t-1}, \qquad p_t^* = \int_{\mathcal{I}_f} p_{jt}^* \, dj \tag{3.2}$$

That is, the (log) aggregate price level at time t is a weighted average of the average price set by resetters and the average price set by non-resetters, p_{t-1}.

Optimal Price Setting A firm re-optimizing in period t will choose the price P_{jt}^* that maximizes the current market value of the profits generated while the price remains effective. Formally,

$$P_{jt}^* = \arg\max_{P_{jt}} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_{jt} \left\{ \Lambda_{t,t+k} \frac{P_{jt} Y_{j,t+k} - W_{t+k} N_{j,t+k}}{P_{t+k}} \right\}$$

where $\Lambda_{t,t+k} \equiv \beta^k \left(\frac{C_{t+k}}{C_t}\right)^{-\sigma}$ is the stochastic discount factor and $\mathbb{E}_{jt}(\cdot)$ denotes firm j's expectation conditional on *its* information set at time t, and subject to the sequence of demand schedules $Y_{j,t+k} = \left(\frac{P_{jt}}{P_{t+k}}\right)^{-\epsilon} C_{t+k}$ and their production technology (3.1). I assume that prices are set before wages. Log-linearizing the resulting first-order condition around the zero inflation steady-state, I obtain the familiar price-setting rule

$$p_{jt}^* = (1-\beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt} \left(p_{t+k} + \Theta \widehat{\mathrm{mc}}_{t+k} \right) \tag{3.3}$$

where $\widehat{\mathrm{mc}}_t = \mathrm{mc}_t - \mathrm{mc}$ is the deviation between real marginal costs and steady-state marginal costs and $\Theta = \frac{1-\alpha}{1-\alpha+\alpha\epsilon}$. Comparing the price-setting rule (3.3) arising in this framework with the one in the benchmark, the only difference comes from the expectation operator. In the benchmark case, information sets are homogeneous and all firms (allowed to act) set the same price. Instead, in this framework, each firm will set a different price based on its own belief structure.

Equilibrium Market clearing in the goods and labor market implies that $c_t = y_t = (1-\alpha)n_t$. Using the equilibrium aggregate labor supply condition, one can write marginal costs

in terms of output, $\text{mc}_t = w_t - p_t = \left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right) y_t$, where σ is the elasticity of intertemporal substitution and φ is the inverse Frisch elasticity. Rewriting output in terms of its gap with respect to the flexible-prices equilibrium,

$$p^*_{jt} = (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\left[p_{t+k} + \Theta\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)\tilde{y}_{t+k}\right] \tag{3.4}$$

which one can rewrite recursively as

$$p^*_{jt} = (1-\beta\theta)\mathbb{E}_{jt}p_t + \frac{\kappa\theta}{1-\theta}\mathbb{E}_{jt}\tilde{y}_t + \beta\theta\mathbb{E}_{jt}p^*_{j,t+1} \tag{3.5}$$

where $\kappa = \frac{(1-\theta)(1-\beta\theta)}{\theta}\Theta\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)$. Condition (3.5) is actually quite intuitive: when a firm j sets its price, it considers how competitive will its price be compared to the average price in the economy (playing a game of strategic complementarities with other firms), which will be the aggregate demand in the economy, and the future conditions since its price will be effective for an unknown number of periods.

Demand side The demand side behaves as in the standard framework. Output gap dynamics are described by the standard DIS curve (3.6), where current output gap depends negatively on the expected real interest rate and positively on future aggregate demand; and nominal interest rates are set by the central bank following a Taylor rule (3.7), in which the central bank reacts to excessive inflation and output by reducing the nominal interest rates, and releases a monetary policy shock (3.8) that has an AR(1) structure:

$$\tilde{y}_t = -\frac{1}{\sigma}\left(i_t - \mathbb{E}_t\pi_{t+1}\right) + \mathbb{E}_t\tilde{y}_{t+1} \tag{3.6}$$

$$i_t = \phi_\pi\pi_t + \phi_y\tilde{y}_t + v_t \tag{3.7}$$

$$v_t = \rho v_{t-1} + \sigma_\varepsilon\varepsilon^v_t, \qquad \varepsilon^v_t \sim \mathcal{N}(0,1) \tag{3.8}$$

The monetary policy shock v_t will be a key object in this economy. It is the only aggregate state variable, and I will assume that firms will have imperfect information on the central bank's action v_t, consistent with the evidence on the transparency policy change by the Fed.

Aggregate Phillips curve In order to derive the aggregate Phillips curve, one can aggregate condition (3.4) across firms.[27] The aggregate Phillips curve can then be written as

$$\pi_t = \kappa\theta \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}_t^f \tilde{y}_{t+k} + (1-\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}_t^f \pi_{t+k} + \left(\overline{\mathbb{E}}_t^f p_{t-1} - p_{t-1} \right) \tag{3.9}$$

where $\pi_t = p_t - p_{t-1}$ is the inflation rate and $\overline{\mathbb{E}}_t^f(\cdot) = \int_{\mathcal{I}_f} \mathbb{E}_{jt}(\cdot)\, dj$ is the average firm expectation operator. Compared to the standard framework, there is an additional term on the right-hand side, the result of firms not perfectly observing the previous price index. Angeletos and Huo (2021) eliminate this term by assuming that firms know the aggregate price level at time $t-1$, but do not extract any information from it.[28] In order to maintain internal consistence in the theoretical framework, I do not make any such assumption.

At this point, it is important to stress that in order to derive condition (3.5) I have not yet specified an information structure. Therefore, the price-setting condition (3.5) and the aggregate Phillips curve (3.9) should be interpreted as a *general* individual price-setting condition and a *general* aggregate Phillips curve.[29]

Information Structure In order to generate heterogeneous beliefs and sticky forecasts, I assume that the information is incomplete and dispersed. Each firm j observes a noisy signal x_{jt} that contains information on the monetary shock v_t, and takes the standard functional form of "outcome plus noise". Formally, signal x_{jt} is described as

$$x_{jt} = v_t + \sigma_u u_{jt}, \qquad \text{with } u_{jt} \sim \mathcal{N}(0,1) \tag{3.10}$$

where signals are agent-specific. This implies that each agent's information set is different, and therefore generates heterogeneous information sets across the population of firms.

An equilibrium must therefore satisfy the individual-level optimal pricing policy functions (3.5), the aggregate DIS curve (3.6), the Taylor rule (3.7), and rational expectation formation

27. We subtract p_{t-1} on both sides, and $\pm\overline{\mathbb{E}}_t^f p_{t-1}$ on the right-hand side.

28. Vives and Yang (2016) motivate this through bounded rationality and inattention, while Angeletos and Huo (2021) argue that inflation contains little statistical information about real variables. Huo and Pedroni (2021) allow for endogenous information, but such a choice complicates the dynamics and the concept of persistence becomes less clear.

29. In the FIRE NK model, agents perfectly observe inflation and output, and face a symmetric Nash equilibrium game, and thus every firm acts as a representative agent firm. In such a case, the individual price-setting curve (3.5) can be aggregated to the well-known New Keynesian Phillips curve (2.4).

should be consistent with the exogenous monetary shock process (3.8) and the signal process (3.10).

Solution Algorithm Here I outline the solution algorithm, and the interested reader is referred to the Proof of Proposition 1 in Appendix A. I first guess that the dynamics of the output gap are endogenous to the aggregate price index and the monetary shock: $\tilde{y}_t = a_y p_{t-1} + b_y p_{t-2} + c_y v_t$ for some unknown coefficients (a_y, b_y, c_y). This allows me to write the individual price-setting condition (3.5) as a beauty contest in which each firm's decision will depend on its own expectation of the fundamental and others' actions. I then compute the expectations. For example, using the Kalman filter, one can write the expectation process as[30]

$$\begin{aligned}\mathbb{E}_{jt}\boldsymbol{Z}_t &= \boldsymbol{\Lambda}\mathbb{E}_{j,t-1}\boldsymbol{Z}_{t-1} + \boldsymbol{K}x_{jt} \\ &= (\boldsymbol{I} - \boldsymbol{\Lambda}L)^{-1}\boldsymbol{K}x_{jt} \\ &= \tilde{\boldsymbol{\Lambda}}(L)x_{jt}, \qquad \boldsymbol{Z}_t = \begin{bmatrix} v_t & p_t & \tilde{y}_t \end{bmatrix}'\end{aligned} \tag{3.11}$$

where I have made use of the lag operator L, and $\tilde{\boldsymbol{\Lambda}}(z) = (\boldsymbol{I} - \boldsymbol{\Lambda}L)^{-1}\boldsymbol{K}$ is a polynomial matrix that depends on the guessed dynamics and the information noise σ_u. I then insert these objects into firm j's price policy function (3.5), and obtain aggregate price dynamics. Finally, I verify the initial guess by introducing the implied price dynamics into the DIS curve (3.6).

Notice that extending the benchmark framework to noisy and dispersed information generates anchoring through expectations, which now follow an autorregressive process. This additional anchoring will result in inflation being more persistent in the noisy information framework, compared to the benchmark setting.

The following proposition outlines inflation and output gap dynamics.

Proposition 1. *Under noisy information the output gap and price level dynamics are given by*

$$\tilde{y}_t = a_y p_{t-1} + b_y p_{t-2} + c_y v_t \tag{3.12}$$

$$p_t = (\vartheta_1 + \vartheta_2)p_{t-1} - \vartheta_1\vartheta_2 p_{t-2} - \psi_\pi \chi_\pi(\vartheta_1, \vartheta_2)v_t \tag{3.13}$$

30. In the case of the Kalman filter, we also need to guess the dynamics of the price level.

where $a_y(\vartheta_1, \vartheta_2)$, $b_y(\vartheta_1, \vartheta_2)$ and $c_y(\vartheta_1, \vartheta_2)$, and where ϑ_1 and ϑ_2 are the reciprocal of the two outside roots of the quartic polynomial

$$\begin{aligned}\mathcal{P}(z) = &-(\beta\theta - z)(1-\theta z)(z-\rho)\,(1-\rho z)\\ &-\tau z\Bigg[(\beta\theta - z)(1-\theta z) + z(1-\theta)(1-\beta\theta)\\ &+ z^2\kappa\theta\frac{\vartheta_1[\sigma(1-\vartheta_2)+\phi_y](\vartheta_1+\vartheta_2-1-\phi_\pi)+(1-\vartheta_2)(\phi_\pi-\vartheta_2)(\sigma+\phi_y)}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\\ &+ z^3\kappa\theta\frac{\vartheta_1\vartheta_2[\sigma(1-\vartheta_1)(1-\vartheta_2)-(\vartheta_1+\vartheta_2-1-\phi_\pi)\phi_y]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\Bigg]\end{aligned}$$

and χ_π is a scalar endogenous to information frictions, with $\tau = \sigma_\varepsilon^2/\sigma_u^2$.

Proof. See Appendix A □

First differencing the price level dynamics (3.13), one can obtain the implied inflation dynamics as

$$\pi_t = (\vartheta_1 + \vartheta_2)\pi_{t-1} - \vartheta_1\vartheta_2\pi_{t-2} - \psi_\pi\chi_\pi(\vartheta_1, \vartheta_2)\Delta v_t \tag{3.14}$$

In the noisy information framework, inflation is intrinsically persistent and its persistence is governed by the new information-related parameters ϑ_1 and ϑ_2, as opposed to the benchmark framework in which it is only extrinsically persistent. The intuition for this result is simple: inflation is partially determined by expectations (see condition (3.9) under noisy information, or (2.4) under complete information). Under noisy information, expectations are anchored and follow an autoregressive process (see (3.11)), which creates the additional source of anchoring in inflation dynamics, measured by the information-related parameters ϑ_1 and ϑ_2.

In the next section I relate the theoretical findings on inflation dynamics to empirical evidence on information frictions, and their fall in the recent decades.

3.2 Calibrating Information Frictions

In the theoretical framework I rationalize the average forecast underreaction through anchoring to priors. A positive β_{rev} will therefore generate intrinsic persistence in inflation dynamics. Yet, this is not enough to explain the *change* in inflation persistence over time. I documented a structural break in belief formation in Section 2.2. This break coincides with a change in the US Federal Reserve's communication policy, which became more transparent

and informative after the mid 1980s. Using survey data on US firms' forecasts, I document a significant sluggishness in responses to new information until the mid 1980s, but no evidence of sluggishness afterwards. In this section, I calibrate the information friction parameter σ_u to match the observed sluggishness in forecasts across time. As argued before, the signal noise became more precise in the dispersed-information model lens. In the next proposition, I relate the previous empirical findings on expectations to model-implied inflation persistence.

Proposition 2. *The theoretical counterpart of the coefficient β_{rev} in (2.2) is given by*

$$\begin{aligned}\beta_{rev} &= \frac{\mathbb{C}\left(forecast\ error_t, revision_t\right)}{\mathbb{V}\left(revision_t\right)} \\ &= \frac{\lambda^3\rho(1-\vartheta_1\lambda)(1-\vartheta_2\lambda)}{(1-\lambda^4)(\rho-\lambda)}\left\{\lambda\frac{\prod_{j=1}^4(\lambda-\xi_j)}{\prod_{k=1}^2(\lambda-\vartheta_k)} - \frac{1-\lambda^2}{\vartheta_1-\vartheta_2}\sum_{k=1}^{2}\frac{\vartheta_k\prod_{j=1}^4(\vartheta_k-\xi_j)}{(1-\lambda\vartheta_k)(\lambda-\vartheta_k)}\right\} \quad (3.15)\end{aligned}$$

where λ is the inside root of the quadratic polynomial $\mathcal{Q}_1(z) = (1-\rho z)(z-\rho) + \frac{\sigma_\varepsilon^2}{\sigma_u^2}z$, and $(\xi_1, \xi_2, \xi_3, \xi_4)$ are the reciprocals of the roots of the quartic polynomial $\mathcal{Q}_2(z) = \phi_0 + \phi_1 z + \phi_2 z^2 + \phi_3 z^3 + \phi_4 z^4$, where $\phi_0 = -\psi_\pi\chi_\pi$, $\phi_1 = \left(\frac{1}{\lambda} - \frac{1}{\rho}\right)\phi_0$, $\phi_2 = \frac{(\rho-\lambda)\phi_0}{\lambda^2\rho}$, $\phi_3 = \frac{(\rho-\lambda)\phi_0[\lambda^3-\vartheta_1-\vartheta_2+\lambda\vartheta_1\vartheta_2]}{\lambda^2\rho(1-\lambda\vartheta_1)(1-\lambda\vartheta_2)}$, and $\phi_4 = \frac{-\lambda^3+\lambda^4\vartheta_2+\lambda^4\vartheta_1-\vartheta_1\vartheta_2[\lambda-(1-\lambda^4)\rho]}{\lambda^2\rho(1-\lambda\vartheta_1)(1-\lambda\vartheta_2)}$.

Proof. See Appendix A. □

The empirical results reported in section 2.2 support a fall in information frictions in recent decades. Proposition 2 maps the theoretical information friction, σ_u, with the Coibion and Gorodnichenko (2015a) estimate. It introduces the model-implied β_{rev} coefficient, which depends on the monetary policy shock persistence ρ and on the information-related parameters ϑ_1, ϑ_2 and λ, where λ, in turn, depends on the persistence parameter and the signal-to-noise ratio. In the noisy information framework, β_{rev} is strictly positive and increases with the degree of information frictions. I show this graphically in Figure 5. In the model lens, this underrevision is the consequence of individual anchoring to priors, and generates forecast underreaction at the aggregate level.

As a last remark, notice that the dynamics generated by the noisy information model (3.14) resemble those generated by the ad-hoc backward-looking models presented in Appendix D.3. However, differently from those ad-hoc frameworks, in the noisy information framework, intrinsic persistence is the result of the micro-founded anchoring in expectations. Extending the model to accommodate noisy information introduces anchoring *through* expectations, for which I have provided empirical evidence, rather than the more ad-hoc con-

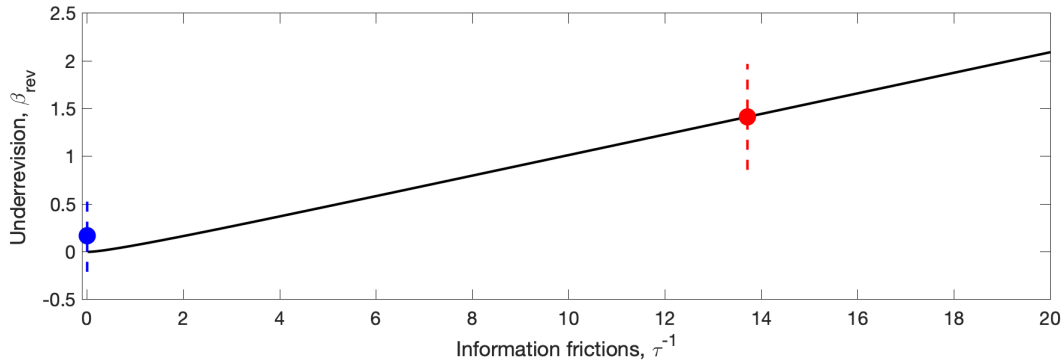


Figure 5: β_{rev} and information frictions τ^{-1}. In red, estimated underrevision coefficient (with 95% confidence interval, dashed line) before 1985. In blue, estimated underrevision coefficient (with 95% confidence interval, dashed line) after 1985.

sumption external habits or price indexation assumptions, for which there is little or no evidence.[31]

4 Results

4.1 Inflation Persistence

In the noisy information framework, inflation persistence is governed by ϑ_1 and ϑ_2. To prove this formally, one can write the inflation first-order autocorrelation as

$$\rho_1 = \frac{(1+\rho)(\vartheta_1+\vartheta_2)+(1-\rho)(\vartheta_1\vartheta_2-1)}{1+\rho\vartheta_1\vartheta_2},$$

which is increasing in both ϑ_1 and ϑ_2. Since the ultimate goal is to understand the break in inflation persistence documented in Section 2.1, the following proposition exposes the determinants of ϑ_1 and ϑ_2, and provides analytical comparative statics.

Proposition 3. *The persistence parameters are*

(i) $\vartheta_1 \in (0, \rho)$

(ii) ϑ_1 *is increasing in* σ_u

31. Havranek et al. (2017) present a meta-analysis of the different estimates of habits in the macro literature and the available micro-estimates. In general, macro models take an index of habits of 0.75, whereas micro-estimates suggest a value around 0.4. On the other hand, the price-indexation model suggests that every price is changed in every period, which is inconsistent with the micro-data estimates provided by Nakamura and Steinsson (2008).

(iii) $\vartheta_2 \in (\theta, 1)$

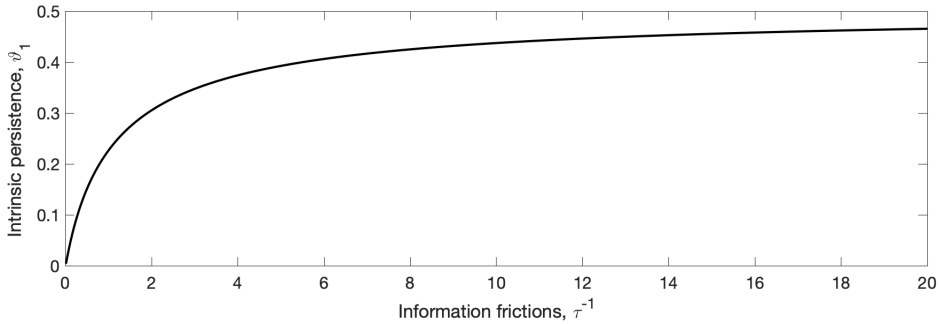
(iv) ϑ_2 *is decreasing in* σ_u

Proof. See Appendix A. □

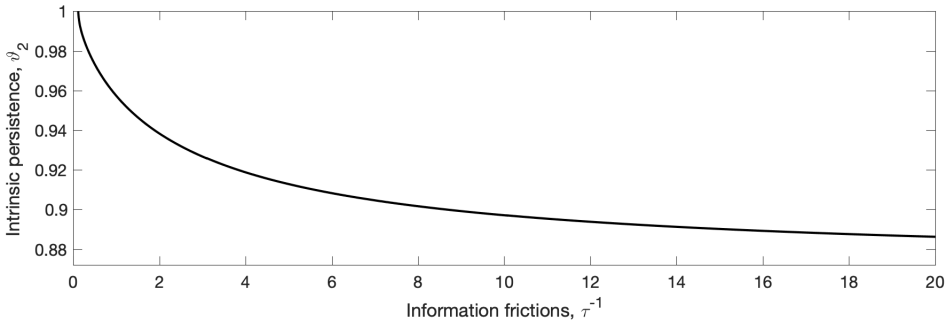
Inflation persistence and information frictions are related through ϑ_1 and ϑ_2. The above proposition is key to understanding the time-varying properties of inflation persistence. First, part (i) establishes that ϑ_1 is bounded by 0 and ρ. Part (ii) states that ϑ_1 is increasing in the degree of information frictions, formalized via the noise of the signal innovation σ_u. A decrease in information frictions reduces inflation first-order autocorrelation through a de-anchoring of individual inflation expectations, which would in turn de-anchor inflation dynamics. Figure 6a plots the level of intrinsic persistence ϑ_1 for different degrees of information frictions, measured by τ^{-1}. Part (iii) establishes that ϑ_2 is bounded by θ and 1. Part (iv) states that ϑ_2 is decreasing in the degree of information frictions. A decrease in information frictions increases inflation first-order autocorrelation through an anchoring of individual inflation expectations, which would in turn anchor inflation dynamics. Figure 6b plots the level of intrinsic persistence ϑ_2 for different degrees of information frictions. In the limit of no information frictions $\sigma_u \to 0$, $\vartheta_1 \to 0$ and $\vartheta_2 \to 1$.

Information frictions do, therefore, have opposing effects on persistence. On the one hand, information frictions lead to an additional persistence through an increase in ϑ_1, the standard mechanism in Angeletos and Huo (2021). On the other hand, there is an additional component ϑ_2 that is decreasing in information frictions. This element arises from the fact that I am solving the NK model in prices, instead of inflation as in Angeletos and Huo (2021) or as in the benchmark setting in Galí (2015) in which prices follow a unit root. Since price dynamics follow (3.2), when firm j forecasts the aggregate price level p_t, she needs to forecast the average action by other firms p_t^*, but also backcast the aggregate price level in the past p_{t-1}. Information frictions relax the forward-lookingness of the model equations, as formalized by Gabaix (2020) and Angeletos and Huo (2021), resulting in price dynamics no longer following a unit root. In the frictionless limit, prices follow a unit root, formalized by $\vartheta_2 \to 1$. However, as shown in Figure 6c, the net result of an increase in information frictions is an increase in the first-order autocorrelation. These key results, coupled with the next result introduced in Proposition 2, will explain the overall fall in inflation persistence.

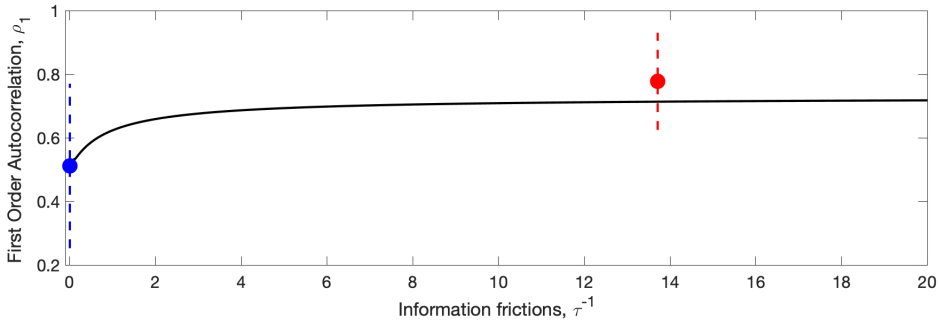
The key finding is that β_{rev} and ρ_1, the theoretical counterparts of Coibion and Gorodnichenko (2015a) underreaction estimate β_{rev} and inflation persistence, are closely related as I show in Figure 6d. The fall in the first-order autocorrelation can be explained by a fall



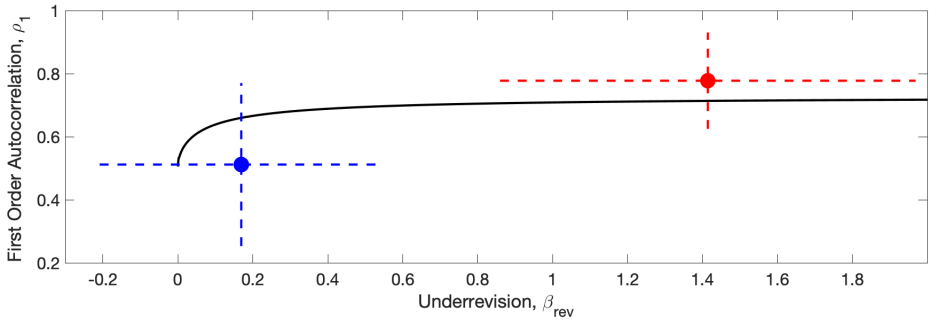
(a) Intrinsic persistence ϑ_1 and information frictions τ^{-1}



(b) Intrinsic persistence ϑ_2 and information frictions τ^{-1}



(c) First-Order Autocorrelation ρ_1 and information frictions τ^{-1}



(d) First-Order Autocorrelation ρ_1 and information frictions β_{rev}

Figure 6: Comparative statics. In red, estimated first-order autocorrelation and underrevision coefficients (with 95% confidence interval, dashed line) before 1985. In blue, estimated first-order autocorrelation and underrevision coefficients (with 95% confidence interval, dashed line) after 1985.

in information frictions. For this quantitative analysis, I use a standard parameterization in the literature, with the only exception of $\theta = 0.872$, which is calibrated to match a Phillips curve slope $\kappa = 0.06$, and $\phi_y = 0.5$ which guarantees the existence of a unique equilibrium as $\sigma_u \to 0$. Finally, I calibrate $\tau = 0.069$ in the pre-1985 sample to match the empirical evidence on β_{rev} in Table III.[32]

Propositions 1-2 establish a direct relation between the first-order autocorrelation of inflation ρ_1 and β_{rev}, our empirical measure of information frictions. Figure 6d shows graphically the monotonically increasing relation between inflation persistence and β_{rev}. In the initial pre-1985 period, with $\beta_{\text{rev}} = 1.501$ from table III, the model-implied inflation first-order autocorrelation is $\rho_1 = 0.716$. In the post-1985 period, with no information frictions, the first-order autocorrelation falls to $\rho_1 = \rho$, which is the persistence of the monetary policy shock in the benchmark framework (see Galí 2015). Comparing our model results to the empirical analysis in Tables I and II, I find that the noisy information framework produces persistence dynamics that lie within the 95% confidence interval, and can explain around 90% of the fall in the point estimate. Noisy information produces such fall in a micro-consistent manner, compared to the more ad-hoc NK models studied in Appendix D.

Role of Calvo Friction In this framework, the first-order autocorrelation of inflation depends on the degree of information frictions, summarized by the two roots ϑ_1 and ϑ_2. A key parameter affecting the transmission of information frictions to the economy is the Calvo innaction probability θ, since it regulates the degree of strategic complementarities on firms' actions. To see this, insert the aggregate price dynamics (3.2) into firm j's best response (3.5),

$$p^*_{jt} = (1-\beta\theta)(1-\theta)\sum_{k=0}^{\infty}\theta^k \mathbb{E}_{jt} p^*_{t-k} + \frac{\kappa\theta}{1-\theta}\mathbb{E}_{jt}\tilde{y}_t + \beta\theta\mathbb{E}_{jt}p^*_{j,t+1}. \tag{4.1}$$

An increase in the Calvo innaction probability has opposing effects on the degree of strategic complementarities within firms. On the one hand, an increase in θ reduces the impact of expected past aggregate actions through a smaller coefficient $(1-\beta\theta)(1-\theta)$. On the other hand, it increases the memory of past expectations on today's actions, $\sum_{k=0}^{\infty}\theta^k\mathbb{E}_{jt}p^*_{t-k}$. It seems natural, however, that this second effect dominates when we look at inflation persistence. Formally, we showed in proposition 3 that ϑ_2 is bounded from below by θ and that it is decreasing in the degree of information frictions σ_u. Therefore, a high θ limits the

32. All parameters are set to the values reported in Table OA.1.

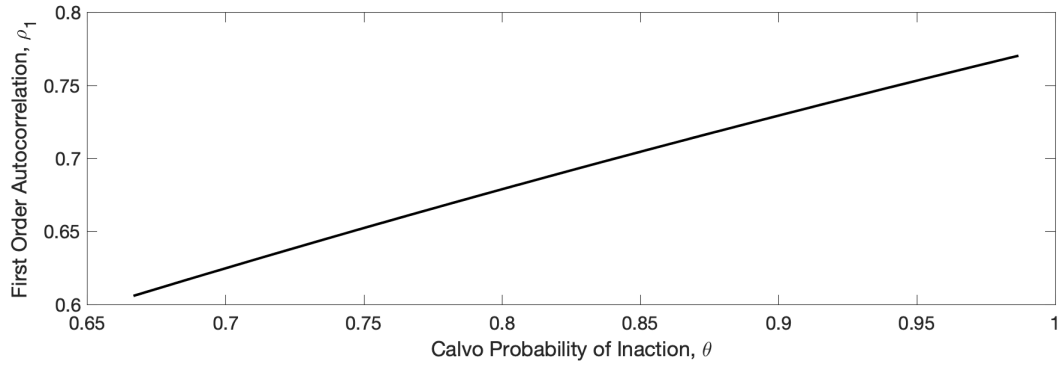


Figure 7: First-order autocorrelation ρ_1 and price friction θ

sensitivity of the root ϑ_2 to changes in σ_u, and helps in generating the desired increase in the first-order autocorrelation after an increase in σ_u, given that ρ_1 is increasing in ϑ_2.

The calibration of the Calvo pricing friction implies a mean price duration of 7.8 quarters. This estimate is in the upper range in the micro literature, although well aligned with the macro literature. Bils and Klenow (2004), Klenow and Kryvtsov (2008), Nakamura and Steinsson (2008), and Goldberg and Hellerstein (2009) find a median price duration of 4.5-11 months in US micro data. Galí (2015) sets $\theta = 0.75$ to match an implied duration of 1 year. Christiano et al. (2011) set $\theta = 0.85$. Auclert et al. (2020) and Afsar et al. (2021) estimate θ between 0.88 and 0.93 from macro data, implying a price duration of 12-14 quarters.

In Figure 7, I plot the implied first-order autocorrelation for different values of the Calvo price friction in the range of the literature. Depending on this parameter, the noisy information framework explains between 40% and 100% of the fall in the point estimate in the first-order autocorrelation.

4.2 The Phillips Curve

In this section I argue that the mainstream finding that the slope of the Phillips curve has fallen in the recent period is simply the result of a misspecified Phillips curve equation (2.4). The derivation of the Phillips curve relies on the FIRE assumption (and implicitly on the Law of Iterated Expectations), for which I find a strong rejection in the data. I then conduct two main exercises. First, in a more theoretical exercise, I use the noisy information framework to rewrite its inflation dynamics as an *as if* FIRE setting with wedges (Angeletos and Huo 2021). According to my theory, the Phillips curve (2.4) needs to be extended with a backward-looking inflation term and significant myopia towards future inflation in the pre-1985 sample period. Once these additional terms are controlled for, and I estimate

a Phillips curve closer to the hybrid version implied by price-indexation settings, I do not find any evidence of a change in κ. Second, by relaxing the FIRE assumption but without any belief structure restriction, the Phillips curve is instead given by (3.9). Instead of replacing expectations of future inflation by its realization, as the literature generally does when estimating condition (2.4), I use the survey forecasts to estimate (3.9) and I do not find any evidence of a change in the slope.

The Wedge Phillips Curve Next, I argue that once I consider a micro-founded Phillips curve that takes into account noisy information, there is no evidence of a change in the slope of the Phillips curve.

Let us first recall inflation dynamics in the standard model. In the benchmark NK model, the Phillips curve is given by (2.4), the DIS curve is given by (3.6), the Taylor rule is given by (3.7) and the monetary policy shock process is given by (3.8). Inserting the Taylor rule (3.7) into the DIS curve (3.6), one can write the model as a system of two first-order stochastic difference equations with reduced-form dynamics $\boldsymbol{x}_t = \boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\varphi}v_t$, where $\boldsymbol{x}_t = [\tilde{y}_t \quad \pi_t \quad p_t]'$ is a 3×1 vector containing output, inflation and prices, and

$$\boldsymbol{\delta} = \frac{1}{\sigma + \phi_y + \kappa\phi_\pi}\begin{bmatrix} \sigma & 1-\beta\phi_\pi & 0 \\ \sigma\kappa & \kappa + \beta(\sigma+\phi_y) & 0 \\ 0 & -1 & 1 \end{bmatrix}, \quad \boldsymbol{\varphi} = \frac{1}{\sigma + \phi_y + \kappa\phi_\pi}\begin{bmatrix} -1 \\ -\kappa \\ 0 \end{bmatrix}.$$

Angeletos and Huo (2021) show that, using the noisy information dynamics (3.12)-(3.14), one can reverse engineer an *as if* system dynamics that mimics the dynamics of the noisy information model, such that the following ad-hoc system of equations

$$\boldsymbol{x}_t = \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} + \boldsymbol{\omega}_f\boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\varphi}v_t \tag{4.2}$$

satisfies the model dynamics for some pair of 3×3 matrices $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$. The next proposition states that, under a certain pair $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$, the ad-hoc economy produces the same dynamics of the noisy information framework.

Proposition 4. *The ad-hoc hybrid dynamics* (4.2) *produces identical dynamics to the noisy information model if* $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$ *satisfy*

$$\begin{aligned} B - \boldsymbol{\varphi} &= \boldsymbol{\omega}_f\boldsymbol{\delta}(AB + \rho B) \\ \boldsymbol{\omega}_b &= (I_3 - \boldsymbol{\omega}_f\boldsymbol{\delta}A)A \end{aligned} \tag{4.3}$$

where

$$A = \begin{bmatrix} 0 & b_y & a_y + b_y \\ 0 & \vartheta_1\vartheta_2 & -(1-\vartheta_1)(1-\vartheta_2) \\ 0 & \vartheta_1\vartheta_2 & \vartheta_1 + \vartheta_2 - \vartheta_1\vartheta_2 \end{bmatrix}, \quad B = \begin{bmatrix} -\psi_y\chi_y(\vartheta_1,\vartheta_2) \\ -\psi_\pi\chi_\pi(\vartheta_1,\vartheta_2) \\ -\psi_\pi\chi_\pi(\vartheta_1,\vartheta_2) \end{bmatrix}$$

$$a_y = \frac{\vartheta_1[\sigma(1-\vartheta_2)+\phi_y](\vartheta_1+\vartheta_2-1-\phi_\pi) + (1-\vartheta_2)(\phi_\pi-\vartheta_2)(\sigma+\phi_y)}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}$$

$$b_y = \frac{\vartheta_1\vartheta_2[\sigma(1-\vartheta_1)(1-\vartheta_2) - (\vartheta_1+\vartheta_2-1-\phi_\pi)\phi_y]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}$$

In particular, the "as if" FIRE Phillips curve dynamics are described by

$$\pi_t = \omega_\pi\pi_{t-1} + \omega_p p_{t-1} + \gamma_y\kappa\tilde{y}_t + \delta_y\mathbb{E}_t\tilde{y}_{t+1} + \delta_\pi\beta\mathbb{E}_t\pi_{t+1} \tag{4.4}$$

where $(\omega_\pi, \omega_p, \gamma_y, \delta_y, \delta_\pi)$ *depend on the* $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$ *pair.*

Proof. See Appendix A. □

The FIRE wedge Phillips curve (4.4), together with a wedge IS curve derived in Appendix A, produces identical dynamics to the noisy information setup derived in section 3. Notice that, in order to derive similar dynamics in a FIRE setup, the Phillips curve needs to be extended with anchoring towards past inflation and myopia towards future, on top of two additional terms on lagged prices and expected output gap. Since all new terms depend on the degree of information frictions σ_u, the model predicts that changes in beliefs will affect the dynamics of the Phillips curve. In particular, the model predicts that as information frictions vanish, i.e., in the benchmark NK model with no information frictions, we have $\omega_{b,11} = \omega_{b,12} = \omega_{b,21} = \omega_{b,22} = \omega_{f,12} = \omega_{f,21} = 0$ and $\omega_{f,11} = \omega_{f,22} = 1$. As a result, $\omega_\pi = \omega_p = \delta_y = 0$, $\gamma_y = 1$, $\delta_\pi = 1$ and the Phillips curve is reduced to the *only*-forward-looking (2.4).

I now test this theoretical prediction in the data by estimating the wedge Phillips curve (4.4), allowing for a structural break in all coefficients after 1985. Following the literature on Phillips curves estimation, I replace expectations of future inflation by realized future inflation and estimate the equation using the generalized method of moments (GMM). In table IV column 1, I report the estimated coefficients for the full sample exercise. I find that the slope of the Phillips curve is small and not significant. In fact, I find that only inflation-related coefficients are significant, suggesting support for backward-lookingness and significant myopia (coefficient well below 0.99). I report the structural break results in

columns 2 and 3. In column 2 I only allow for a structural break on the contemporaneous output gap coefficient. I find no evidence of a structural break in the slope (i.e., no evidence of flattening in the Phillips curve). In column 3 I explore if there has been any other structural break in the dynamics of the Phillips curve. I find a structural break in lagged and forward inflation: in recent decades the Phillips curve has become *more* forward-looking and less backward-looking. This last result aligns well with the documented drop in inflation persistence and information frictions, and with the mechanism dynamics proposed by the noisy information framework.

Notice that condition (4.3) does not uniquely determine the set of weights $\boldsymbol{\omega}_f$ that is consistent with the noisy information dynamics. Different weights in $\boldsymbol{\omega}_f$ are consistent with noisy information dynamics, although the dynamics are unique.[33] I explore which set of wedges $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$ is consistent with the documented dynamics and with the findings in table IV. Since I do not find any evidence of the relevance of the lagged price level and forward output gap, I choose wedges such that they produce the well-known hybrid Phillips curve. The following corollary provides us with the hybrid wedge Phillips curve.

Corollary 1. *The hybrid Phillips curve*

$$\pi_t = \omega_\pi \pi_{t-1} + \kappa \tilde{y}_t + \delta_\pi \beta \mathbb{E}_t \pi_{t+1} + \chi v_t \tag{4.5}$$

produces identical dynamics to the "as if" FIRE Phillips curve (4.4)*, where* $(\omega_\pi, \delta_\pi, \chi)$ *depend on the* $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$ *pair. As information frictions vanish,* $\omega_\pi = \chi = 0$ *and* $\delta_\pi = 1$*.*

Proof. See Appendix A. □

First, notice that the *as if* model produces anchoring, which lines up with the strong inflation persistence during that period, and myopia towards future inflation. As before, the noisy information model suggests that anchoring and myopia in the hybrid Phillips curve (4.5) should vanish in the post-1985 sample. Estimating the micro-founded hybrid NK Phillips curve (4.5), reported in table IV (columns 4 and 5), I find that one cannot reject the null that, since the structural break in 1985:Q1, (*i*) anchoring has gone to zero and (*ii*) myopia has disappeared. In the light of these noisy estimates, I take the "Phillips curve flattening puzzle" as a result of misspecification in the standard Phillips curve.

33. Intuitively, agents' actions can be anchored/myopic with respect to aggregate output or inflation.

	(1) Wedge Phillips Curve	(2) Break Output	(3) Break All	(4) Hybrid Phillips Curve	(5) Break Hybrid
π_{t-1}	0.447^{***}	0.496^{***}	0.738^{***}	0.453^{***}	0.720^{***}
	(0.0978)	(0.110)	(0.165)	(0.0891)	(0.131)
$\pi_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$			-0.671^{**}		-0.597^{**}
			(0.267)		(0.232)
p_{t-1}	0.00354	−0.0164	0.0363		
	(0.0162)	(0.0243)	(0.202)		
$p_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$			−0.105		
			(0.256)		
$\tilde{y}_t$	0.0633	0.134	0.378	−0.000813	0.0566
	(0.112)	(0.123)	(0.240)	(0.0166)	(0.0488)
$\tilde{y}_t \times \mathbb{1}_{\{t \geq t^*\}}$		−0.0726	−0.127		−0.0143
		(0.0632)	(0.291)		(0.0781)
$\tilde{y}_{t+1}$	−0.0695	−0.104	−0.343		
	(0.126)	(0.125)	(0.212)		
$\tilde{y}_{t+1} \times \mathbb{1}_{\{t \geq t^*\}}$			0.0963		
			(0.328)		
π_{t+1}	0.540^{***}	0.509^{***}	0.226	0.539^{***}	0.273^{**}
	(0.102)	(0.108)	(0.215)	(0.0885)	(0.129)
$\pi_{t+1} \times \mathbb{1}_{\{t \geq t^*\}}$			0.866^{***}		0.643^{***}
			(0.335)		(0.244)
Observations	202	202	202	202	202

HAC (1 lag) robust standard errors in parentheses

Instrument set: four lags of effective federal funds rate, CBO Output gap, GDP Deflator growth rate, Commodity Inflation, M2 growth rate, spread between long- and short-run interest rate and labor share.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IV: Estimates of regression (4.4)

Controlling for Imperfect Expectations In order to obtain the results on inflation persistence, I have assumed a particular belief structure, rational expectations but noisy and dispersed information. In this section I take a step back and instead take an agnostic stance on expectation formation. Let us start from the aggregate Phillips curve (3.9), derived under mild assumptions on beliefs.[34] Inflation is now related to current *and future* output through two different channels: the slope of the Phillips curve, κ, and firms' expectation formation process. In order to test for a potential structural break in the slope *controlling for* non-standard expectations, I regress the *general* Phillips curve (3.9) (truncated at $k = 4$), for which I do not assume a particular information structure, using real GDP and GDP Deflator growth forecast data from the SPF. I set β and θ to their quarterly values 0.99 and 0.872, and regress

$$\pi_t = \alpha_1 + \alpha_2 \tilde{y}_t^e + \alpha_3 \pi_t^e + \eta_t \tag{4.6}$$

where $\eta_t = \left(\overline{\mathbb{E}}_t^f p_{t-1} - p_{t-1}\right) +$ truncation error, $\tilde{y}_t^e = \theta \sum_{k=0}^{4} (\beta\theta)^k \overline{\mathbb{E}}_t^f \tilde{y}_{t+k}$ and $\pi_t^e = \sum_{k=0}^{4} (\beta\theta)^k \overline{\mathbb{E}}_t^f \pi_{t+k}$ denote the truncated sums of expected real GDP and inflation. I use standard GMM methods by instrumenting for expectations with 4-quarter lagged annual inflation and real GDP growth expectations. The results are reported in table V. In column 1, I report the full sample coefficients. I find that κ is small, consistent with the choice of κ and similar to the value found by Hazell et al. (2020). In column 2, I regress its (output) structural break version. This is the only specification that suggests a structural break on the slope. However, when I also consider a potential structural break in inflation, I find an estimate of κ that aligns well with the model assumption, and I find no evidence of a structural break in the Phillips curve slope.[35]

Summary To sum up, I find that once I control for imperfect expectations and a potential change in their dynamics, I do not find any evidence of a structural break in the slope of the Phillips curve. First, I showed that the noisy information model can explain the change in the dynamics of the Phillips curve as a reshuffle between backward-lookingness and forward-lookingness via changes in belief formation. Second, I documented empirically that controlling for non-standard expectations, proxied by the forecasts submitted by professional forecasters, I do not find any evidence of a change in the slope of the Phillips curve.

34. I have only assumed that the law of iterated expectations holds at the individual level.
35. I repeat the analysis using the Livingston Survey on Appendix B, and find similar results.

	(1) Full Sample	(2) Break Output	(3) Break All
$\tilde{y}_t^e$	-0.00692	0.132^{**}	0.0909^{*}
	(0.0177)	(0.0515)	(0.0531)
$\tilde{y}_t^e \times \mathbb{1}_{\{t \geq t^*\}}$		-0.103^{***}	0.0112
		(0.0356)	(0.0649)
π_t^e	0.262^{***}	0.214^{***}	0.237^{***}
	(0.0121)	(0.0223)	(0.0240)
$\pi_t^e \times \mathbb{1}_{\{t \geq t^*\}}$			-0.0932^{**}
			(0.0402)
Observations	199	199	199

HAC robust standard errors in parentheses (1 lag)
Instrument set: four lags of forecasts of annual real GDP growth and annual GDP Deflator growth
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table V: Estimates of regression (4.6)

5 Conclusion

In this paper I explain the fall in inflation persistence since the mid 1980s through changes in beliefs. State-of-the-art monetary models face significant challenges in explaining this fall in inflation persistence. I show that, by relaxing the FIRE assumption in the benchmark NK framework, the model is able to generate the documented fall in persistence. Using micro-data on inflation expectations from the Survey of Professional Forecasters (SPF), I argue that agents became more informed about inflation after the change in the Federal Reserve disclosure policy, which endogenously lowers the intrinsic persistence in inflation dynamics.

I revisit different theories that produce a structural relation between inflation and other forces in the economy. I show that a variety of NK models cannot explain the fall in inflation persistence. Since the benchmark model is purely forward-looking, inflation exhibits no intrinsic persistence, and its dynamic properties are now inherited from monetary policy shocks. However, I document that the persistence of monetary policy shocks has not changed over time. Acknowledging that purely forward-looking models cannot generate anchoring or intrinsic persistence, I extend the benchmark model to incorporate a backward-looking dimension. I show that the change in the monetary stance now affects inflation intrinsic persistence. The effect is small, however. Then, I show that the noisy and dispersed information

extension is consistent with the micro-data evidence on belief formation, and generates anchoring or intrinsic inflation persistence. Using SPF data, I document that a structural break in expectation formation, resulting in agents being more informed about inflation, is contemporaneous to the fall in inflation persistence. The model can therefore explain the fall in inflation persistence in a micro-consistent manner.

I discuss the consequences of noisy and dispersed information on the dynamics of the Phillips curve, and the lack of flattening. In the noisy information model, the Phillips curve is enlarged with anchoring and myopia. Consistent with the theory, I find that both anchoring and myopia vanish after the reduction of information frictions in the mid 1980s. Finally, taking an agnostic stance on expectations, I show that there is any evidence of a change in the Phillips curve slope once I control for imperfect expectations.

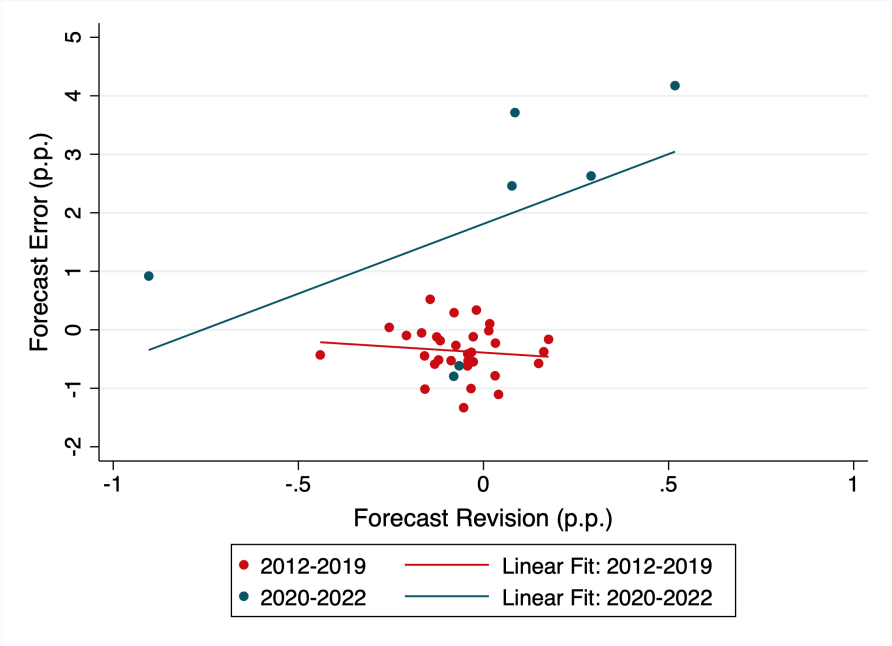
Will the 2020-2022 inflation be persistent? In this paper I have only considered data up until 2020:Q1. The evidence provided points towards a lessening in the underrevision behavior of agents and a fall in inflation persistence since the mid-1980s. Taking these results together would make the reader conclude that current inflation will only be temporary (or, at least, less persistent than before the mid-1980s). However, having a look at the 2020:Q1-2022:Q2 data, one could argue that the underrevision behavior (see figure 8a) and inflation persistence (see figure 8b) are stricking back.[36] Although admittedly speculative, these findings suggest that central banks should focus on their communication in the coming quarters if they want to reduce the current inflation persistence. This theory is imperfect, however, since it abstains from cost-push shocks and the bottlenecks arising from input-output network of the economy. This suggests avenues for follow-up research, in which belief formation frictions interact with the input-output structure of the economy.

References

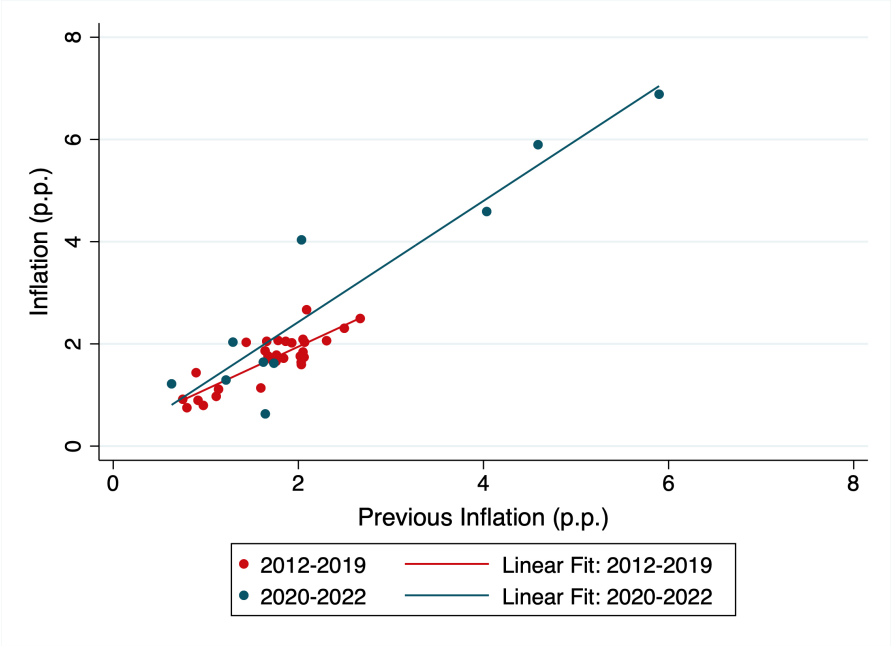
Afsar, Atahan, Jose-Elias Gallegos, Richard Jaimes, and Edgar Silgado. 2021. "Reconciling Empirics and Theory: The Behavioral Hybrid New Keynesian Model."

Allen, Franklin, Stephen Morris, and Hyun Song Shin. 2006. "Beauty contests and iterated expectations in asset markets." *Review of Financial Studies.* ISSN: 08939454.

36. I include the 2012-2019 period to show that the increase in information frictions and persistence is not driven by forward guidance, a major policy change that could, in principle, have dampened agents' understanding of monetary policy due to its novelty.



(a) Scatter plot of ex-ante average forecast error (vertical axis) and average forecast revisions (horizontal axis), computed using the SPF and vintage GDP Deflator data. Red dots correspond to 2012-2019 observations, and blue dots correspond to observations after 2019.



(b) Scatter plot contemporaneous inflation (vertical axis) and one-quarter lagged inflation (horizontal axis). Red dots correspond to 2012-2019 observations, and blue dots correspond to observations after 2019.

Figure 8: Scatter plots of forecast underrevision and inflation's first-order autocorrelation.

Angeletos, George-Marios, and Zhen Huo. 2021. "Myopia and Anchoring." *American Economic Review* 111 (4): 1166–1200.

Angeletos, George-marios, Zhen; Huo, and Karthik A. Sastry. 2020. "Imperfect Macroeconomic Expectations: Theory and Evidence." *Working Paper.*

Ascari, Guido, and Luca Fosso. 2021. "The Inflation Rate Disconnect Puzzle: On the International Component of Trend Inflation and the Flattening of the Phillips Curve." *Working Paper.*

Ascari, Guido, and Argia M. Sbordone. 2014. *The macroeconomics of trend inflation.*

Atkeson, Andrew, and Lee E. Ohanian. 2001. "Are Phillips curves useful for forecasting inflation?" *Quarterly Review* 25 (Win): 2–11.

Auclert, Adrian, Matthew Rognlie, and Ludwig Straub. 2020. "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model." *Working Paper.*

Bacchetta, Philippe, and Eric Van Wincoop. 2006. "Can information heterogeneity explain the exchange rate determination puzzle?" *American Economic Review.* ISSN: 00028282.

Ball, Laurence, and Sandeep Mazumder. 2011. "Inflation Dynamics and the Great Recession." *Brookings Papers on Economic Activity* 42 (1 (Spring): 337–405.

Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. "Are technology improvements contractionary?" *American Economic Review.* ISSN: 00028282.

Benati, Luca, and Paolo Surico. 2008. "Evolving U.S. Monetary Policy and the Decline of Inflation Predictability." *Journal of the European Economic Association* 6 (2/3): 634–646. ISSN: 15424766, 15424774.

Bernanke, Ben S., Thomas Laubach, Frederic S. Mishkin, and Adam S. Posen. 1999. *Inflation Targeting: Lessons from the International Experience.* Princeton University Press.

Bils, Mark, and Peter J. Klenow. 2004. "Some evidence on the importance of sticky prices." *Journal of Political Economy.* ISSN: 00223808.

Blinder, Alan S. 2000. "Central-Bank Credibility: Why Do We Care? How Do We Build It?" *American Economic Review* 90 (5): 1421–1431.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. *Over-reaction in Macroeconomic Expectations.* Technical report.

Broer, Tobias, and Alexandre N Kohlhas. 2019. "Forecaster (Mis-)Behavior."

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45. ISSN: 0022-3808. arXiv: arXiv:1011.1669v3.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. "When Is the Government Spending Multiplier Large?" *Journal of Political Economy* 119 (1): 78–121. ISSN: 00223808, 1537534X.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. "Monetary policy rules and macroeconomic stability: Evidence and some theory." *Quarterly Journal of Economics.* ISSN: 00335533.

Cogley, Timothy, and Argia M. Sbordone. 2008. "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve." *The American Economic Review* 98 (5): 2101–2126. ISSN: 00028282.

Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent. 2010. "Inflation-Gap Persistence in the US." *American Economic Journal: Macroeconomics* 2 (1): 43–69.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015a. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review.* ISSN: 00028282.

———. 2015b. "Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation." *American Economic Journal: Macroeconomics* 7 (1): 197–232.

———. 2012. "What Can Survey Forecasts Tell Us about Information Rigidities?" *Journal of Political Economy* 120 (1): 116–159. ISSN: 00223808, 1537534X.

Cukierman, Alex, and Allan H. Meltzer. 1986. "A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information." *Econometrica* 54 (5): 1099–1128. ISSN: 00129682, 14680262.

del Negro, Marco, Marc P. Giannoni, and Christina Patterson. 2012. "The Forward Guidance Puzzle." *SSRN Electronic Journal.*

del Negro, Marco, Michele Lenza, Giorgio Primiceri, and Andrea Tambalotti. 2020. "What's up with the Phillips Curve?" *National Bureau of Economic Research.* ISSN: 0898-2937.

Faust, Jon, and Lars E. O. Svensson. 2002. "The Equilibrium Degree of Transparency and Control in Monetary Policy." *Journal of Money, Credit and Banking* 34 (2): 520–539. ISSN: 00222879, 15384616.

———. 2001. "Transparency and Credibility: Monetary Policy with Unobservable Goals." *International Economic Review* 42 (2): 369–397. ISSN: 00206598, 14682354.

Fernald, John. 2014. "A quarterly, utilization-adjusted series on total factor productivity."

Francis, Neville, Michael T. Owyang, Jennifer E. Roush, and Riccardo DiCecio. 2014. "A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks." *Review of Economics and Statistics.* ISSN: 15309142.

Fuhrer, Jeff, and George Moore. 1995. "Inflation Persistence." *The Quarterly Journal of Economics* 110 (1): 127–159. ISSN: 00335533, 15314650.

Fuhrer, Jeffrey C. 2010. *Inflation Persistence.*

Gabaix, Xavier. 2020. "A behavioral new Keynesian model." *American Economic Review.* ISSN: 19447981.

Galí, Jordi. 2015. *Monetary policy, inflation, and the business cycle: An introduction to the new Keynesian framework and its applications: Second edition.* ISBN: 9781400866274.

———. 1999. "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?" *American Economic Review.* ISSN: 00028282.

Goldberg, Pinelopi K., and Rebecca Hellerstein. 2009. *How rigid are producer prices?* Technical report.

Goldstein, Nathan, and Yuriy Gorodnichenko. 2019. "Macroeconomic Persistence and Forward Information: What Else Can We Learn from Survey Expectations?" *Working Paper.*

Goodfriend, Marvin. 1986. "Monetary mystique: Secrecy and central banking." *Journal of Monetary Economics* 17 (1): 63–92. ISSN: 0304-3932.

Grossman, Sanford J., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *The American Economic Review* 70 (3): 393–408. ISSN: 00028282.

Hall, Robert E. 2011. "The Long Slump." *American Economic Review* 101 (2): 431–69.

Hamilton, James Douglas. 1994. *Time Series Analysis.* 1–799. Princeton University Press. ISBN: 0-691-04289-6.

Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit formation in consumption: A meta-analysis." *European Economic Review.* ISSN: 00142921.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson. 2020. *The Slope of the Phillips Curve: Evidence from U.S. States.* NBER Working Papers 28005. National Bureau of Economic Research, Inc, October.

Holston, Kathryn, Thomas Laubach, and John C. Williams. 2017. "Measuring the natural rate of interest: International trends and determinants." *Journal of International Economics.* ISSN: 18730353.

Huo, Zhen, and Marcelo Pedroni. 2021. "Dynamic Information Aggregation: Learning from the Past."

Huo, Zhen, and Naoki Takayama. 2018. "Rational Expectations Models with Higher Order Beliefs."

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2011. "Investment shocks and the relative price of investment." *Review of Economic Dynamics.* ISSN: 10942025.

Klenow, Peter J., and Oleksiy Kryvtsov. 2008. "State-Dependent or Time-Dependent Pricing: Does it Matter for Recent U.S. Inflation?*." *The Quarterly Journal of Economics* 123, no. 3 (August): 863–904. ISSN: 0033-5533. eprint: https://academic.oup.com/qje/article-pdf/123/3/863/5434839/123-3-863.pdf.

Kydland, Finn E., and Edward C. Prescott. 1982. "Time to Build and Aggregate Fluctuations." *Econometrica.* ISSN: 00129682.

Lindé, Jesper, and Mathias Trabandt. 2019. *Resolving the Missing Deflation Puzzle.* CEPR Discussion Papers 13690. C.E.P.R. Discussion Papers, April.

Lindsey, David E. 2003. *A Modern History of FOMC Communication: 1975-2002.*

Lorenzoni, Guido. 2009. "A theory of demand shocks." *American Economic Review.* ISSN: 00028282.

Lubik, Thomas A., and Frank Schorfheide. 2003. "Computing sunspot equilibria in linear rational expectations models." *Journal of Economic Dynamics and Control* 28 (2): 273–285.

———. 2004. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *American Economic Review* 94 (1): 190–217.

Lucas, Robert E. 1972. "Expectations and the neutrality of money." *Journal of Economic Theory.* ISSN: 10957235.

McLeay, Michael, and Silvana Tenreyro. 2020. "Optimal Inflation and the Identification of the Phillips Curve." *NBER Macroeconomics Annual* 34:199–255. eprint: https://doi.org/10.1086/707181.

Morris, Stephen, and Hyun Song Shin. 2002. "Social value of public information." *American Economic Review.* ISSN: 00028282.

Morris, Stephen, Hyun Song Shin, Nobuhiro Kiyotaki, Marios Angeletos, and Andrew Atkeson. 2006. "Inertia of forward-looking expectations." In *American Economic Review.*

Nakajima, Jouchi. 2011. "Time-varying parameter VAR model with stochastic volatility: An overview of methodology and empirical applications." *Monetary and Economic Studies* 29:107–142.

Nakamura, Emi, and Jón Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models." *Quarterly Journal of Economics* 123 (4): 1415–1464. ISSN: 0033-5533.

Nekarda, Cj, and Va Ramey. 2010. "The cyclical behavior of the price-cost markup." *. . . paper, UCSD and Federal Reserve Board, . . .*

Nimark, Kristoffer. 2008. "Dynamic pricing and imperfect common knowledge." *Journal of Monetary Economics.* ISSN: 03043932.

Ramey, V. A. 2016. "Macroeconomic Shocks and Their Propagation." In *Handbook of Macroeconomics.* ISBN: 9780444594877.

Romer, Christina D, and David H Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review* 94 (4): 1055–1084. ISSN: 0002-8282.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2005. "Optimal Inflation Stabilization in a Medium-Scale Macroeconomic Model." *NBER Working Paper No. 11854*. ISSN: 1098-6596.

Smets, Frank, and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review*. ISSN: 00028282.

Solow, Robert M. 1957. "Technical Change and the Aggregate Production Function." *The Review of Economics and Statistics*. ISSN: 00346535.

Stock, James H., and Mark W. Watson. 2019. *Slack and Cyclically Sensitive Inflation*. NBER Working Papers 25987. National Bureau of Economic Research, Inc, June.

———. 2007. "Why has U.S. inflation become harder to forecast?" In *Journal of Money, Credit and Banking*.

Vives, Xavier, and Liyan Yang. 2016. "A Model of Costly Interpretation of Asset Prices." *SSRN Electronic Journal* (January).

Woodford, Michael. 2001. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Working Paper*.

———. 2003a. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Information and Expectations in Modern Macroeconomics*.

———. 2003b. *Interest and prices: foundations of a theory of monetary policy*. Princeton, NJ: Princeton Univ. Press. ISBN: 0691010498.

Appendix

A Proofs of Propositions in Main Text

Proof of Proposition 1. Under noisy information in the firm side, the individual price policy functions are given by (3.5). Let us guess that the equilibrium output gap dynamics will take the form of

$$\tilde{y}_t = a_y p_{t-1} + b_y p_{t-2} + c_y v_t \tag{A.1}$$

Making use of the guess I can rewrite the price-setting condition as

$$p^*_{it} = \frac{\kappa\theta c_y}{1-\theta}\mathbb{E}_{it}v_t + \frac{\kappa\theta b_y}{1-\theta}\mathbb{E}_{it}p_{t-2} + \frac{\kappa\theta a_y}{1-\theta}\mathbb{E}_{it}p_{t-1} + (1-\beta\theta)\mathbb{E}_{it}p_t + \beta\theta\mathbb{E}_{it}p^*_{i,t+1} \tag{A.2}$$

We now turn to solving the expectation terms in (A.2). We can write the fundamental representation of the signal process as a system containing (3.8) and (3.10), which admits the following state-space representation

$$\begin{aligned} \boldsymbol{Z}_t &= \boldsymbol{F}\boldsymbol{Z}_{t-1} + \boldsymbol{\Phi}\boldsymbol{s}_{it} \\ \boldsymbol{X}_{it} &= \boldsymbol{H}\boldsymbol{Z}_t + \boldsymbol{\Psi}\boldsymbol{s}_{it} \end{aligned} \tag{A.3}$$

with $\boldsymbol{F} = \rho$, $\boldsymbol{\Phi} = \begin{bmatrix}\sigma_\varepsilon & 0\end{bmatrix}$, $\boldsymbol{Z}_t = v_t$, $\boldsymbol{s}_{it} = \begin{bmatrix}\varepsilon^v_t \\ u_{it}\end{bmatrix}$, $\boldsymbol{H} = 1$, $\boldsymbol{\Psi} = \begin{bmatrix}0 & \sigma_u\end{bmatrix}$ and $\boldsymbol{X}_{it} = x_{it}$. It is convenient to rewrite the uncertainty parameters in terms of precision: define $\tau_\varepsilon \equiv \frac{1}{\sigma^2_\varepsilon}$ and $\tau_u \equiv \frac{1}{\sigma^2_u}$. The signal system can be written as

$$\boldsymbol{X}_{it} = \frac{\sigma_\varepsilon}{1-\rho L}\varepsilon^v_t + \sigma_u u_{it} = \begin{bmatrix}\frac{\tau_\varepsilon^{-1/2}}{1-\rho L} & \tau_u^{-1/2}\end{bmatrix}\begin{bmatrix}\varepsilon^v_t \\ u_{it}\end{bmatrix} = \boldsymbol{M}(L)\boldsymbol{s}_{it}, \qquad \boldsymbol{s}_{it} \sim \mathcal{N}(0, I) \tag{A.4}$$

The Wold theorem states that there exists another representation of the signal process (A.4),

$$\boldsymbol{X}_{it} = \boldsymbol{B}(L)\boldsymbol{w}_{it}$$

such that $\boldsymbol{B}(z)$ is invertible and $\boldsymbol{w}_{it} \sim (0, \boldsymbol{V})$ is white noise. Hence, we can write the following equivalence

$$\boldsymbol{X}_{it} = \boldsymbol{M}(L)\boldsymbol{s}_{it} = \boldsymbol{B}(L)\boldsymbol{w}_{it} \tag{A.5}$$

In the Wold representation of $\boldsymbol{X}_{it}$, observing $\{\boldsymbol{X}_{it}\}$ is equivalent to observing $\{\boldsymbol{w}_{it}\}$, and $\{\boldsymbol{X}_i^t\}$ and $\{\boldsymbol{w}_i^t\}$ contain the same information. Furthermore, note that the Wold representation has the property that, using the equivalence (A.5), both processes share the autocovariance generating function

$$\rho_{xx}(z) = \boldsymbol{M}(z)\boldsymbol{M}'(z^{-1}) = \boldsymbol{B}(z)\boldsymbol{V}\boldsymbol{B}'(z^{-1})$$

Given the state-space representation of the signal process (A.26), optimal expectations of the exogenous fundamental take the form of a Kalman filter

$$\mathbb{E}_{it}v_t = (I - \boldsymbol{K}\boldsymbol{H})\boldsymbol{F}\mathbb{E}_{it-1}v_{t-1} + \boldsymbol{K}x_{it} = \lambda\mathbb{E}_{it-1}v_{t-1} + \boldsymbol{K}x_{it}$$

where $\boldsymbol{K}$ is given by

$$\boldsymbol{K} = \boldsymbol{P}\boldsymbol{H}'\boldsymbol{V}^{-1} \tag{A.6}$$

$$\boldsymbol{P} = \boldsymbol{F}[\boldsymbol{P} - \boldsymbol{P}\boldsymbol{H}'\boldsymbol{V}^{-1}\boldsymbol{H}\boldsymbol{P}]\boldsymbol{F} + \boldsymbol{\Phi}\boldsymbol{\Phi}' \tag{A.7}$$

We still need to find the unknowns $\boldsymbol{B}(z)$ and $\boldsymbol{V}$. Propositions 13.1-13.4 in Hamilton (1994) provide us with these objects. Unknowns $\boldsymbol{B}(z)$ and $\boldsymbol{V}$ satisfy

$$\boldsymbol{B}(z) = I + \boldsymbol{H}(I - \boldsymbol{F}z)^{-1}\boldsymbol{F}\boldsymbol{K}$$
$$\boldsymbol{V} = \boldsymbol{H}\boldsymbol{P}\boldsymbol{H}' + \boldsymbol{\Psi}\boldsymbol{\Psi}'$$

I can write (A.7) as

$$\boldsymbol{P}^2 + \boldsymbol{P}[(1-\rho^2)\sigma_u^2 - \sigma_\varepsilon^2] - \sigma_\varepsilon^2\sigma_u^2 = 0 \tag{A.8}$$

from which we can infer that $\boldsymbol{P}$ is a scalar. Denote $k = \boldsymbol{P}^{-1}$ and rewrite (A.8) as

$$\sigma_u^2\sigma_\varepsilon^2 k^2 = [(1-\rho^2)\sigma_u^2 - \sigma_\varepsilon^2]k + 1 \implies k = \frac{\tau_\varepsilon}{2}\left\{1 - \rho^2 - \frac{\tau_u}{\tau_\varepsilon} \pm \sqrt{\left[\frac{\tau_u}{\tau_\varepsilon} - (1-\rho^2)\right]^2 + 4\frac{\tau_u}{\tau_\varepsilon}}\right\}$$

I also need to find $\boldsymbol{K}$. Now that we have found $\boldsymbol{P}$ in terms of model primitives, we can obtain $\boldsymbol{K}$ using condition (A.6)

$$\boldsymbol{K} = \frac{1}{1 + k\sigma_u^2}$$

We can finally write λ as

$$\lambda = (I - \boldsymbol{KH})\boldsymbol{F} = \frac{k\sigma_u^2\rho}{1+k\sigma_u^2} = \frac{1}{2}\left[\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\varepsilon} \pm \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\varepsilon}\right)^2 - 4}\right] \tag{A.9}$$

One can show that one of the roots $\lambda_{1,2}$ lies inside the unit circle and the other lies outside as long as $\rho \in (0,1)$, which guarantees that the Kalman expectation process is stationary and unique. We set λ to the root that lies inside the unit circle (the one with the '$-$' sign). Notice that I can also write $\boldsymbol{V}$ in terms of λ

$$\boldsymbol{V} = k^{-1} + \sigma_u^2 = \frac{\rho}{\lambda\tau_u}$$

where I have used the identity $k = {}^{\lambda\tau_u}/_{(\rho\,-\,\lambda)}$. Finally, I can obtain $\boldsymbol{B}(z)$

$$\boldsymbol{B}(z) = 1 + \frac{\rho z}{(1-\rho z)(1+k\sigma_u^2)} = \frac{1-\lambda z}{1-\rho z}$$

and therefore one can verify that

$$\boldsymbol{B}(z)\boldsymbol{V}\boldsymbol{B}'(z^{-1}) = \boldsymbol{M}(z)\boldsymbol{M}'(z^{-1})$$
$$\frac{\rho}{\lambda\tau_u}\frac{(1-\lambda z)(z-\lambda)}{(1-\rho z)(z-\rho)} = \frac{\tau_\varepsilon z}{(1-\rho z)(z-\rho)} + \tau_u$$

Let us now move to the forecast of *endogenous* variables. Consider a variable $f_t = A(L)\boldsymbol{s}_{it}$. Applying the Wiener-Hopf prediction filter, we can obtain the forecast as

$$\mathbb{E}_{it} f_t = \left[A(z)\boldsymbol{M}'(z^{-1})\boldsymbol{B}(z^{-1})^{-1}\right]_+ \boldsymbol{V}^{-1}\boldsymbol{B}(z)^{-1}x_{it}$$

where $[\cdot]_+$ denotes the annihilator operator.[37]

Recall from condition (A.2) that we are interested in obtaining $\mathbb{E}_{jt}v_t$, $\mathbb{E}_{jt}\pi_t$ and $\mathbb{E}_{jt}\pi_{j,t+1}$. Just as we did in the example above, we need to find the $A(z)$ polynomial for each of the forecasted variables. Let us start from the exogenous fundamental v_t to verify that the Kalman and Wiener-Hopf filters result in the same forecast. I can write the fundamental as

$$v_t = \begin{bmatrix} \frac{\tau_\varepsilon^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \boldsymbol{s}_{it} = A_v(L)\boldsymbol{s}_{it}$$

37. See Online Appendix H for more details on the Wiener-Hopf prediction filter and the annihilator operator.

Let us now move to the endogenous variables. In this case we need to guess (and verify) that each agent i's policy function takes the following form[38]

$$p_{it} = h(L)x_{it}$$

Aggregate price level can then be expressed as

$$p_t = (1-\theta)\int h(L)x_{it}\,di + \theta p_{t-1} = (1-\theta)h(L)\frac{\tau_\varepsilon^{-1/2}}{(1-\rho L)(1-\theta L)}\varepsilon_t^v$$

Using the guesses, I have

$$\begin{aligned} p_{t-k} &= \begin{bmatrix}(1-\theta)\tau_\varepsilon^{-1/2}\frac{h(L)L^k}{(1-\rho L)(1-\theta L)} & 0\end{bmatrix}\boldsymbol{s}_{it} = A_{pk}(L)\boldsymbol{s}_{it} \\ p_{i,t+1} &= \frac{h(L)}{L}\boldsymbol{M}(L)\boldsymbol{s}_{it} = \begin{bmatrix}\tau_\varepsilon^{-1/2}\frac{h(L)}{L(1-\rho L)} & \tau_u^{-1/2}\frac{h(L)}{L}\end{bmatrix}\boldsymbol{s}_{it} = A_i(L)\boldsymbol{s}_{it}\end{aligned}$$

We are now armed with the necessary objects in order to obtain the three different forecasts,

$$\begin{aligned} \mathbb{E}_{it}v_t &= \left[A_v(z)\boldsymbol{M}'(z^{-1})\boldsymbol{B}(z^{-1})^{-1}\right]_+\boldsymbol{V}^{-1}\boldsymbol{B}(z)^{-1}x_{it} \\ &= \left[\begin{bmatrix}\frac{\tau_\varepsilon^{-1/2}}{1-\rho z} & 0\end{bmatrix}\begin{bmatrix}\frac{z\tau_\varepsilon^{-1/2}}{z-\rho} \\ \tau_u^{-1/2}\end{bmatrix}\frac{z-\rho}{z-\lambda}\right]_+\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\ &= \left[\frac{z}{\tau_\varepsilon(1-\rho z)(z-\lambda)}\right]_+\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\ &= \left[\frac{\phi_v(z)}{z-\lambda}\right]_+\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\ &= \frac{\phi_v(z)-\phi_v(\lambda)}{z-\lambda}\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\ &= \frac{\lambda\tau_u}{\rho\tau_\varepsilon(1-\rho\lambda)}\frac{1}{1-\lambda z}x_{it} = G_1(z)x_{it} \end{aligned} \tag{A.10}$$

$$\begin{aligned} \mathbb{E}_{it}p_{t-k} &= \left[A_{pk}(z)\boldsymbol{M}'(z^{-1})\boldsymbol{B}(z^{-1})^{-1}\right]_+\boldsymbol{V}^{-1}\boldsymbol{B}(z)^{-1}x_{it} \\ &= \left[\begin{bmatrix}(1-\theta)\tau_\varepsilon^{-1/2}\frac{h(z)z^k}{(1-\rho z)(1-\theta z)} & 0\end{bmatrix}\begin{bmatrix}\frac{z\tau_\varepsilon^{-1/2}}{z-\rho} \\ \tau_u^{-1/2}\end{bmatrix}\frac{z-\rho}{z-\lambda}\right]_+\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \end{aligned}$$

38. In this framework agents only observe signals. As a result, the policy function can only depend on current and past signals.

$$
\begin{aligned}
&= \left[\frac{h(z)z^{k+1}}{(1-\rho z)(z-\lambda)(1-\theta z)}\right]_+ \frac{(1-\theta)\lambda\tau_u}{\tau_\varepsilon\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \left[\frac{\phi_\pi(z)}{z-\lambda}\right]_+ \frac{(1-\theta)\lambda\tau_u}{\tau_\varepsilon\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \frac{\phi_\pi(z)-\phi_\pi(\lambda)}{z-\lambda}\frac{(1-\theta)\lambda\tau_u}{\rho\tau_\varepsilon}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= (1-\theta)\frac{\lambda\tau_u}{\rho\tau_\varepsilon}\left[\frac{h(z)z^{k+1}}{1-\theta z} - h(\lambda)\lambda^{k+1}\frac{1-\rho z}{(1-\rho\lambda)(1-\theta\lambda)}\right]\frac{1}{(1-\lambda z)(z-\lambda)}x_{it} = G_2(z)x_{it}
\end{aligned} \tag{A.11}
$$

$$
\begin{aligned}
\mathbb{E}_{it}p_{i,t+1} &= \left[A_i(z)\boldsymbol{M}'(z^{-1})\boldsymbol{B}(z^{-1})^{-1}\right]_+ \boldsymbol{V}^{-1}\boldsymbol{B}(z)^{-1}x_{it} \\
&= \left[\begin{bmatrix}\tau_\varepsilon^{-1/2}\frac{h(z)}{z(1-\rho z)} & \tau_u^{-1/2}\frac{h(z)}{z}\end{bmatrix}\begin{bmatrix}\frac{z\tau_\varepsilon^{-1/2}}{z-\rho} \\ \tau_u^{-1/2}\end{bmatrix}\frac{z-\rho}{z-\lambda}\right]_+ \frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \left[\frac{h(z)}{\tau_\varepsilon(1-\rho z)(z-\lambda)} + \frac{h(z)(z-\rho)}{\tau_u z(z-\lambda)}\right]_+ \frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \left\{\left[\frac{h(z)}{\tau_\varepsilon(1-\rho z)(z-\lambda)}\right]_+ + \left[\frac{h(z)(z-\rho)}{\tau_u z(z-\lambda)}\right]_+\right\}\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \left\{\left[\frac{\phi_{i,1}(z)}{z-\lambda}\right]_+ + \left[\frac{\phi_{i,2}(z)}{z(z-\lambda)}\right]_+\right\}\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \left\{\frac{\phi_{i,1}(z)-\phi_{i,1}(\lambda)}{z-\lambda} + \frac{\phi_{i,2}(z)-\phi_{i,2}(\lambda)}{\lambda(z-\lambda)} - \frac{\phi_{i,2}(z)-\phi_{i,2}(0)}{\lambda z}\right\}\frac{\lambda\tau_u}{\rho}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= \frac{\lambda}{\rho}\left\{\frac{h(z)}{z-\lambda}\left[\frac{\tau_u}{\tau_\varepsilon(1-\rho z)} + \frac{z-\rho}{z}\right] - \frac{h(\lambda)}{z-\lambda}\left[\frac{\tau_u}{\tau_\varepsilon(1-\rho\lambda)} + \frac{\lambda-\rho}{\lambda}\right] - \frac{\rho h(0)}{\lambda z}\right\}\frac{1-\rho z}{1-\lambda z}x_{it} \\
&= G_3(z)x_{it}
\end{aligned} \tag{A.12}
$$

where

$$
\phi_v(z) = \frac{z}{\tau_\varepsilon(1-\rho z)}, \quad \phi_\pi(z) = \frac{h(z)z}{(1-\rho z)(1-\theta z)}, \quad \phi_{i,1}(z) = \frac{h(z)}{\tau_\varepsilon(1-\rho z)}, \quad \phi_{i,2}(z) = \frac{h(z)(z-\rho)}{\tau_u}
$$

Rearranging terms, we obtain (A.13)-(A.15). We can show that expectations satisfy

$$
\mathbb{E}_{it}v_t = \left(1-\frac{\lambda}{\rho}\right)\frac{1}{1-\lambda L}x_{it} \tag{A.13}
$$

$$
\mathbb{E}_{it}p_{t-k} = (1-\theta)\left(1-\frac{\lambda}{\rho}\right)\left[\frac{h(z)z^{k+1}(1-\rho\lambda)}{1-\theta z} - \frac{h(\lambda)\lambda^{k+1}(1-\rho z)}{1-\theta\lambda}\right]\frac{1}{(1-\lambda z)(z-\lambda)} \tag{A.14}
$$

$$
\mathbb{E}_{it}p^*_{i,t+1} = \left\{\frac{h(z)}{z-\lambda}\left[\left(1-\frac{\lambda}{\rho}\right)\frac{1-\rho\lambda}{1-\rho z} + \frac{\lambda(z-\rho)}{\rho z}\right] - \frac{h(0)}{z}\right\}\frac{1-\rho z}{1-\lambda z} \tag{A.15}
$$

Recall the best response for firm i, condition (A.2). In order to be consistent with firm optimization, the policy function $h(z)$ must satisfy (A.2) at all times and signals. Plugging the obtained expressions and rearranging by $h(z)$, we can write

$$\tilde{C}(z)h(z)x_{it} = d[z; h(\lambda), h(0)]x_{it}$$

where

$$\begin{aligned}
\tilde{C}(z) &= (z-\beta\theta)(1-\theta z)(z-\lambda)(1-\lambda z) \\
&\quad - z^2\kappa\theta\left(\frac{(1-\theta)(1-\beta\theta)}{\kappa\theta} + za_y + z^2 b_y\right)\left(1-\frac{\lambda}{\rho}\right)(1-\rho\lambda) \\
&= \lambda\Bigg\{(\beta\theta - z)(1-\theta z)(z-\rho)\left(z-\frac{1}{\rho}\right) \\
&\quad - \frac{\tau}{\rho}z\left[(\beta\theta - z)(1-\theta z) + \kappa\theta\left(\frac{(1-\theta)(1-\beta\theta)}{\kappa\theta} + za_y + z^2 b_y\right)z\right]\Bigg\} \\
&= \lambda C(z) \\
d[z; h(\lambda), h(0)] &= \frac{\kappa\theta c_y}{1-\theta}\left(1-\frac{\lambda}{\rho}\right)z(z-\lambda)(1-\theta z) \\
&\quad - h(\lambda)\frac{\lambda}{1-\theta\lambda}\left(1-\frac{\lambda}{\rho}\right)\kappa\theta\left(\frac{(1-\theta)(1-\beta\theta)}{\kappa\theta} + \lambda a_y + \lambda^2 b_y\right)z(1-\rho z)(1-\theta z) \\
&\quad - h(0)\beta\theta(1-\rho z)(z-\lambda)(1-\theta z)
\end{aligned}$$

where I have used the following identity from the Kalman filter

$$\lambda + \frac{1}{\lambda} = \rho + \frac{1}{\rho} + \frac{\tau}{\rho} \implies (\rho-\lambda)(1-\rho\lambda) = \lambda\tau$$

Notice that we can write polynomial $\tilde{C}(z)$ in terms of its roots as

$$\tilde{C}(z) = \theta\lambda\left(1-\frac{\tau\kappa b_y}{\rho}\right)(z-\zeta_1)(z-\zeta_2)(z-\vartheta_1^{-1})(z-\vartheta_2^{-1})$$

where ζ_1, ζ_2 are the inside roots of $C(z)$, and ϑ_1 and ϑ_2 are the reciprocals of the outside roots. In order to have a causal $h(z)$ polynomial, we need to eliminate the inside roots in its denominator, $\lambda C(z)$. I choose $h(0)$ and $h(\lambda)$ so that $d[\zeta_1; h(0), h(\lambda)] = 0$ and $d[\zeta_2; h(0), h(\lambda)] = 0$.

As a result, I can write

$$d[z; h(0), h(\lambda)] = \frac{\kappa\theta\lambda\tau c_y}{(1-\theta)\rho(1-\rho\zeta_1)(1-\rho\zeta_2)}(z-\zeta_1)(z-\zeta_2)(1-\theta z)$$

and hence the policy function is

$$h(z) = \frac{\kappa c_y}{1-\theta}\frac{\tau\vartheta_1\vartheta_2}{(\rho-\tau\kappa b_y)(1-\rho\zeta_1)(1-\rho\zeta_2)}\frac{1-\theta z}{(1-\vartheta_1 z)(1-\vartheta_2 z)} \tag{A.16}$$

Hence, aggregate price dynamics follow

$$p_t = (1-\theta)\frac{\int h(L)x_{it}\,di}{1-\theta L} = (1-\theta)\frac{h(L)}{1-\theta L}v_t = \kappa c_y\frac{\tau\vartheta_1\vartheta_2}{(\rho-\tau\kappa b_y)(1-\rho\zeta_1)(1-\rho\zeta_2)}\frac{1}{(1-\vartheta_1 L)(1-\vartheta_2 L)}v_t$$

We can therefore write inflation dynamics as

$$\begin{aligned}\pi_t = (1-L)p_t &= \kappa c_y\frac{\tau\vartheta_1\vartheta_2}{(\rho-\tau\kappa b_y)(1-\rho\zeta_1)(1-\rho\zeta_2)}\frac{1}{(1-\vartheta_1 L)(1-\vartheta_2 L)}v_t \\ &= (\vartheta_1+\vartheta_2)\pi_{t-1} - \vartheta_1\vartheta_2\pi_{t-2} + c_p\Delta v_t \end{aligned} \tag{A.17}$$

where $c_p = \kappa c_y\frac{\tau\vartheta_1\vartheta_2}{(\rho-\tau\kappa b_y)(1-\rho\zeta_1)(1-\rho\zeta_2)}$.

Inserting inflation dynamics into the DIS equation (3.6) I can obtain output gap dynamics

$$\begin{aligned}\tilde{y}_t &= \frac{1}{\sigma}(-\phi_\pi p_t + \phi_\pi p_{t-1} + \sigma\mathbb{E}_t\tilde{y}_{t+1} + \mathbb{E}_t p_{t+1} - p_t - v_t) \\ &= \frac{(\sigma a_y + \vartheta - \phi_\pi)(1+\vartheta) + \phi_\pi + \sigma b_y - \vartheta}{\sigma}p_{t-1} - \frac{(\sigma a_y + \vartheta - \phi_\pi)\vartheta}{\sigma}p_{t-2} \\ &\quad - \frac{1-\rho(c_p - \sigma c_y) - (\sigma a_y + \vartheta - \phi_\pi)c_p}{\sigma}v_t \end{aligned} \tag{A.18}$$

In order to be consistent with our earlier guess (A.1), it must be that

$$\begin{aligned} a_y &= \frac{\vartheta_1[\sigma(1-\vartheta_2)+\phi_y](\vartheta_1+\vartheta_2-1-\phi_\pi) + (1-\vartheta_2)(\phi_\pi-\vartheta_2)(\sigma+\phi_y)}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]} \\ b_y &= \frac{\vartheta_1\vartheta_2[\sigma(1-\vartheta_1)(1-\vartheta_2) - (\vartheta_1+\vartheta_2-1-\phi_\pi)\phi_y]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]} \end{aligned}$$

and two additional coefficients (c_p, c_y) irrelevant for persistence.

Finally, we can rewrite the $C(z)$ polynomial as

$$C(z) = \frac{\lambda}{\rho}\Bigg\{(z-\beta\theta)(1-\theta z)(z-\rho)(1-\rho z) - z^3(1-\theta)(1-\beta\theta)(1-\rho)^2$$
$$+ \tau z\left[(1-\theta z)(z-\beta\theta) - z(1-\theta)(1-\beta\theta) + \frac{(1-z)z^2\theta\kappa}{\sigma}\vartheta\right]\Bigg\}$$

$$C(z) = \frac{\lambda}{\rho}\Bigg\{ -(\beta\theta - z)(1-\theta z)(z-\rho)(1-\rho z)$$
$$- \tau z\Bigg[(\beta\theta - z)(1-\theta z) + z(1-\theta)(1-\beta\theta)$$
$$+ z^2\kappa\theta\frac{\vartheta_1[\sigma(1-\vartheta_2)+\phi_y](\vartheta_1+\vartheta_2-1-\phi_\pi) + (1-\vartheta_2)(\phi_\pi-\vartheta_2)(\sigma+\phi_y)}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}$$
$$+ z^3\kappa\theta\frac{\vartheta_1\vartheta_2[\sigma(1-\vartheta_1)(1-\vartheta_2) - (\vartheta_1+\vartheta_2-1-\phi_\pi)\phi_y]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\Bigg]\Bigg\}$$

□

Proof of Proposition 2. We are interested in obtaining $\beta_{rev} = \frac{\mathbb{C}(\text{forecast error}_t, \text{revision}_t)}{\mathbb{V}(\text{revision}_t)}$. Using the results from the proof of Proposition 1 that we can write the forecast error as

$$\begin{aligned}
\pi_{t+3,t} - \overline{\mathbb{E}}_t^f \pi_{t+3,t} &= p_{t+3} - p_{t-1} - \overline{\mathbb{E}}_t^f (p_{t+3} - p_{t-1}) \\
&= \frac{\phi_0 + \phi_1 z + \phi_2 z^2 + \phi_3 z^3 + \phi_4 z^4}{(1-\lambda z)(1-\vartheta_1 z)(1-\vartheta_2 z)}\varepsilon_{t+3}^v \\
&= \phi_0 \frac{(1-\xi_1 z)(1-\xi_2 z)(1-\xi_3 z)(1-\xi_4 z)}{(1-\lambda z)(1-\vartheta_1 z)(1-\vartheta_2 z)}\varepsilon_{t+3}^v \\
&= \frac{\phi_0(\lambda-\xi_1)(\lambda-\xi_2)}{(\lambda-\vartheta_1)(\lambda-\vartheta_2)}\sum_{k=0}^{k}\lambda^k[\varepsilon_{t+3-k}^v - (\xi_3+\xi_4)\varepsilon_{t+2-k}^v + \xi_3\xi_4\varepsilon_{t+1-k}^v] \\
&\quad - \frac{\phi_0(\vartheta_1-\xi_1)(\vartheta_1-\xi_2)}{(\lambda-\vartheta_1)(\vartheta_1-\vartheta_2)}\sum_{k=0}^{k}\vartheta_1^k[\varepsilon_{t+3-k}^v - (\xi_3+\xi_4)\varepsilon_{t+2-k}^v + \xi_3\xi_4\varepsilon_{t+1-k}^v] \\
&\quad + \frac{\phi_0(\vartheta_2-\xi_1)(\vartheta_2-\xi_2)}{(\lambda-\vartheta_2)(\vartheta_1-\vartheta_2)}\sum_{k=0}^{k}\vartheta_2^k[\varepsilon_{t+3-k}^v - (\xi_3+\xi_4)\varepsilon_{t+2-k}^v + \xi_3\xi_4\varepsilon_{t+1-k}^v]
\end{aligned}$$

where $\phi_0 = c_p$, $\phi_1 = \left(\frac{1}{\lambda} - \frac{1}{\rho}\right) c_p$, $\phi_2 = \frac{(\rho-\lambda)c_p}{\lambda^2\rho}$, $\phi_3 = \frac{(\rho-\lambda)c_p[\lambda^3-\vartheta_1-\vartheta_2+\lambda\vartheta_1\vartheta_2]}{\lambda^2\rho(1-\lambda\vartheta_1)(1-\lambda\vartheta_2)}$, $\phi_4 = \frac{-\lambda^3+\lambda^4\vartheta_2+\lambda^4\vartheta_1-\vartheta_1\vartheta_2[\lambda-(1-\lambda^4)\rho]}{\lambda^2\rho(1-\lambda\vartheta_1)(1-\lambda\vartheta_2)}$ and $(\xi_1, \xi_2, \xi_3, \xi_4)$ are the reciprocals of the roots of the polynomial $\phi_0 + \phi_1 z + \phi_2 z^2 + \phi_3 z^3 + \phi_4 z^4$.

The average forecast revision is given by

$$\begin{aligned}
\overline{\mathbb{E}}_t^f \pi_{t+3,t} - \overline{\mathbb{E}}_{t-1}^f \pi_{t+3,t} &= \overline{\mathbb{E}}_t^f (p_{t+3} - p_{t-1}) - \overline{\mathbb{E}}_{t-1}^f (p_{t+3} - p_{t-1}) \\
&\frac{c_p(\rho-\lambda)(1-\lambda^4)}{\rho\lambda^3(1-\vartheta_1\lambda)(1-\vartheta_2\lambda)(1-\lambda z)}\varepsilon_t^v \\
&= \frac{c_p(\rho-\lambda)(1-\lambda^4)}{\rho\lambda^3(1-\vartheta_1\lambda)(1-\vartheta_2\lambda)} \sum_{k=0}^{\infty} \lambda^k \varepsilon_{t-k}^v
\end{aligned}$$

and we can finally write β_{rev} as

$$\begin{aligned}
\beta_{\text{rev}} &= \frac{\mathbb{C}(\text{forecast error}_t, \text{revision}_t)}{\mathbb{V}(\text{revision}_t)} \\
&= \frac{\lambda^3\rho(1-\vartheta_1\lambda)(1-\vartheta_2\lambda)}{(1-\lambda^4)(\rho-\lambda)} \Bigg\{ \frac{\lambda(\lambda-\xi_1)(\lambda-\xi_2)(\lambda-\xi_3)(\lambda-\xi_4)}{(\lambda-\vartheta_1)(\lambda-\vartheta_2)} \\
&\quad - (1-\lambda^2)\left[\frac{\vartheta_1(\vartheta_1-\xi_1)(\vartheta_1-\xi_2)(\vartheta_1-\xi_3)(\vartheta_1-\xi_4)}{(1-\lambda\vartheta_1)(\lambda-\vartheta_1)(\vartheta_1-\vartheta_2)} + \frac{\vartheta_2(\vartheta_2-\xi_1)(\vartheta_2-\xi_2)(\vartheta_2-\xi_3)(\vartheta_2-\xi_4)}{(1-\lambda\vartheta_2)(\lambda-\vartheta_2)(\vartheta_1-\vartheta_2)}\right] \Bigg\}
\end{aligned}$$

□

Proof of Proposition 3. Let us first show that the polynomial described by $C(z)$ has two inside roots and two outside roots. To do so, I evaluate $C(z)$ at $z = \{0, \lambda, 1, \rho^{-1}\}$

$$C(0) = \beta\theta\lambda > 0$$

$$\begin{aligned}
C(\lambda) = -\theta\kappa\lambda^2\left(1-\frac{\lambda}{\rho}\right)(1-\rho\lambda)\Bigg[&\frac{(1-\theta)(1-\beta\theta)}{\theta\kappa} + \frac{\lambda}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]} \\
&\times [\sigma(1-\vartheta_1)(1-\vartheta_2)(\phi_\pi - \vartheta_2 - \vartheta_1(1-\lambda\vartheta_2)) + \phi_y(\vartheta_1(1-\lambda\vartheta_2)(\vartheta_1+\vartheta_2-1-\phi_\pi) + (1-\vartheta_2)(\phi_\pi-\vartheta_2))]\Bigg] < 0
\end{aligned} \tag{A.19}$$

$$C(1) = \frac{\lambda}{\rho}\left\{(1-\theta)(1-\beta\theta)(1-\rho)^2 + \frac{\kappa\theta\tau(1-\vartheta_1)(1-\vartheta_2)[\vartheta_1(\sigma(1-\vartheta_2)+\phi_y) - (\phi_\pi-\vartheta_2)(\sigma+\phi_y)]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\right\} > 0 \tag{A.20}$$

$$\begin{aligned}
C(\rho^{-1}) = -\frac{\theta\lambda\tau}{\rho^5}\Bigg\{&(1-\rho)\rho(1-\rho\beta) + \frac{\kappa}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]} \\
&\times [\sigma(1-\vartheta_1)(1-\vartheta_2)[\vartheta_1(\vartheta_2-\rho)+\rho(\phi_\pi-\vartheta_2)] + \phi_y[\vartheta_1(\vartheta_2-\rho)(1+\phi_\pi-\vartheta_1-\vartheta_2) + \rho(1-\vartheta_2)(\phi_\pi-\vartheta_2)]]\Bigg\} < 0
\end{aligned}$$

Notice that all conditions are trivially satisfied except for the second (A.19) and third (A.20) conditions, which depend on the model parameterization. Combining both conditions, we obtain the restriction

$$\frac{\tau(1-\vartheta_1)(1-\vartheta_2)[(\phi_\pi-\vartheta_2)(\sigma+\phi_y)-\vartheta_1(\sigma(1-\vartheta_2)+\phi_y)]}{(1-\rho)^2[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]} < \frac{(1-\theta)(1-\beta\theta)}{\theta\kappa} <$$
$$< -\frac{\lambda\left[\sigma(1-\vartheta_1)(1-\vartheta_2)\left(\phi_\pi-\vartheta_2-\vartheta_1(1-\lambda\vartheta_2)\right)+\phi_y\left(\vartheta_1(1-\lambda\vartheta_2)(\vartheta_1+\vartheta_2-1-\phi_\pi)+(1-\vartheta_2)(\phi_\pi-\vartheta_2\right.\right.}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}$$

It turns out that a standard calibration satisfies both conditions except for the limit case $\sigma_u = 0$. Hence, I can conclude that the polynomial has two roots inside the unit circle and two roots outside, and all of them are real.

Let us now show that $\vartheta_1 < \rho$. First, it is important to note that λ is the inside root of the polynomial

$$\mathcal{C}(z) = z^2 - \left(\frac{1}{\rho}+\rho+\frac{\tau}{\rho}\right)z+1$$

which has one inside root and one outside root if $\rho < 1$ and $\tau > 0$. Furthermore

$$\mathcal{C}(0) = 1 > 0$$
$$\mathcal{C}(\rho) = -\frac{\tau}{\rho} < 0$$

and, hence, $\lambda < \rho$. We have shown that $C(\rho^{-1}) < 0$, and we have $C(\vartheta_1^{-1}) = 0$. I also know that the function $C(z)$ is always positive for values larger than ϑ_1^{-1}, and hence I can infer $\rho^{-1} < \vartheta_1^{-1}$ and $\vartheta_1 < \rho$. In order to show that $\lambda > \vartheta_1$, I obtain

$$C(\lambda^{-1}) = -\frac{\theta\kappa\tau}{\rho\lambda}\left\{\frac{(1-\theta)(1-\beta\theta)}{\theta}\right.$$
$$\left. + \frac{\sigma(1-\vartheta_1)(1-\vartheta_2)[\vartheta_1(\vartheta_2-\lambda)+\lambda(\phi_\pi-\vartheta_2)]+\phi_y[\vartheta_1(\vartheta_2-\lambda)(1+\phi_\pi-\vartheta_1-\vartheta_2)+\lambda(1-\vartheta_2)(\phi_\pi-\vartheta_2)]}{\lambda^2[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\right\} < 0$$

Following the same argument, knowing that $\lambda < 1$ and that the function $C(z)$ is negative for values of $\vartheta_1^{-1} > z > \vartheta_2^{-1}$, I can write $\lambda^{-1} < \vartheta_1^{-1}$ and $\lambda > \vartheta_1$. Hence I have proved the relation $\vartheta_1 < \lambda < \rho$.

Let us now show that $\theta < \vartheta_2 < 1$. We already proved that $\vartheta_2^{-1} > 1$, which implies that

$\vartheta_2 < 1$. We have that

$$C(\theta^{-1}) = -\frac{\kappa\tau\lambda}{\rho\theta^3}\Bigg\{\frac{(1-\theta)(1-\beta\theta)\theta}{\kappa} + \frac{\sigma(1-\vartheta_1)(1-\vartheta_2)[\vartheta_1(\vartheta_2-\theta)+\theta(\phi_\pi-\vartheta_2)]+\phi_y[\vartheta_1(\vartheta_2-\theta)(1+\phi_\pi-\vartheta_1-\vartheta_2)+\theta(1-\vartheta_2)(\phi_\pi-\vartheta_2)]}{[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]}\Bigg\} < 0$$

Notice that $C(\theta^{-1}) < 0$, given that $\theta < 1$, implies that $\theta^{-1} > \vartheta_2^{-1}$ and delivers the result $\vartheta_2 < \theta$. To sum up, the following relation holds: $0 < \vartheta_1 < \lambda < \rho < \theta < \vartheta_2 < 1$.

Finally, I show that ϑ_1 is increasing in σ_u. First, let us obtain the effect of an increase in τ and ϑ around $C(\vartheta^{-1})$,

$$\begin{aligned}
\frac{\partial C(\vartheta_1^{-1})}{\partial\tau} &= \frac{\theta\lambda(1-\vartheta_1)}{\rho\vartheta_1^3[\sigma(1-\vartheta_1)+\phi_y]}[\vartheta_1(1-\vartheta_1)(1+\beta)\sigma-\kappa(\phi_\pi-\vartheta_1)-\phi_y(1-\beta\vartheta_1)] > 0\\
\frac{\partial C(\vartheta_2^{-1})}{\partial\tau} &= \frac{\theta\lambda(1-\vartheta_2)}{\rho\vartheta_2^3[\sigma(1-\vartheta_2)+\phi_y]}[\vartheta_2(1-\vartheta_2)(1+\beta)\sigma-\kappa(\phi_\pi-\vartheta_2)-\phi_y(1-\beta\vartheta_2)] > 0\\
\frac{\partial C(\vartheta_1^{-1})}{\partial\vartheta_1} &= \frac{\theta\kappa\tau\lambda(\vartheta_2-\vartheta_1)}{\rho\vartheta_1^4[\sigma(1-\vartheta_1)+\phi_y]^2[\sigma(1-\vartheta_2)+\phi_y]}\\
&\quad\times[[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y](1-\vartheta_1)+\phi_y[(\sigma+\phi_y)(\phi_\pi-\vartheta_1-\vartheta_2)+\sigma\vartheta_1\vartheta_2]] > 0\\
\frac{\partial C(\vartheta_2^{-1})}{\partial\vartheta_2} &= -\frac{\theta\kappa\tau\lambda(\vartheta_2-\vartheta_1)}{\rho\vartheta_2^4[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y]^2}\\
&\quad\times[[\sigma(1-\vartheta_1)+\phi_y][\sigma(1-\vartheta_2)+\phi_y](1-\vartheta_2)+\phi_y[(\sigma+\phi_y)(\phi_\pi-\vartheta_1-\vartheta_2)+\sigma\vartheta_1\vartheta_2]] < 0
\end{aligned}$$

Using the Implicit Function Theorem I can infer that $\vartheta_1'(\tau) < 0$ and $\vartheta_2'(\tau) > 0$, and so ϑ_1 (ϑ_2) is increasing (decreasing) in σ_u. □

Proof of Proposition 4. In the benchmark NK model the Phillips curve is given by (2.4), the DIS curve is given by (3.6), the Taylor rule is given by (3.7) and the monetary policy shock process is given by (3.8). Inserting the Taylor rule (3.7) into the DIS curve (3.6), one can write the model as a system of two first-order stochastic difference equations

$$\tilde{A}\boldsymbol{x}_t = \tilde{B}\mathbb{E}_t\boldsymbol{x}_{t+1} + \tilde{C}v_t \tag{A.21}$$

where $\boldsymbol{x}_t = [\tilde{y}_t \quad \pi_t \quad p_t]'$ is a 3×1 vector containing output, inflation and prices, $\tilde{A}$ is a 3×3

coefficient matrix, $\tilde{B}$ is a 3×3 coefficient matrix and $\tilde{C}$ is a 3×1 vector satisfying

$$\tilde{A} = \begin{bmatrix} \sigma + \phi_y & \phi_\pi & 0 \\ -\kappa & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \sigma & 1 & 0 \\ 0 & \beta & 0 \\ 0 & -1 & 1 \end{bmatrix}, \quad \text{and} \quad \tilde{C} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

Premultiplying the system by $\tilde{A}^{-1}$ we obtain $\boldsymbol{x}_t = \boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\varphi}v_t$, where

$$\boldsymbol{\delta} = \tilde{A}^{-1}\tilde{B}, \quad \boldsymbol{\varphi} = \tilde{A}^{-1}\tilde{C}$$

In the dispersed information framework, structural-form dynamics are given by $A_s\boldsymbol{x}_t = B_s\boldsymbol{x}_{t-1} + C_s v_t$ where

$$A_s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B_s = \begin{bmatrix} 0 & b_y & a_y + b_y \\ 0 & 0 & -1 \\ 0 & -b_p & a_p + b_p \end{bmatrix}, \quad \text{and} \quad C_s = \begin{bmatrix} c_y \\ 0 \\ c_p \end{bmatrix}$$

Premultiplying by A_s^{-1} we obtain the reduced-form dynamics $\boldsymbol{x}_t = A\boldsymbol{x}_{t-1} + Bv_t$, where

$$A = A_s^{-1}B_s, \quad B = A_s^{-1}B_s$$

Using the Method for Undetermined Coefficients, the ad-hoc behavioral dynamics and the noisy information dynamics are observationally equivalent if

$$\begin{aligned} A\boldsymbol{x}_{t-1} + Bv_t &= \boldsymbol{\varphi}v_t + \boldsymbol{\omega}_f\boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} \\ &= \boldsymbol{\varphi}v_t + \boldsymbol{\omega}_f\boldsymbol{\delta}\mathbb{E}_t(A\boldsymbol{x}_t + Bv_{t+1}) + \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} \\ &= \boldsymbol{\varphi}v_t + \boldsymbol{\omega}_f\boldsymbol{\delta}(A\boldsymbol{x}_t + B\mathbb{E}_t v_{t+1}) + \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} \\ &= \boldsymbol{\varphi}v_t + \boldsymbol{\omega}_f\boldsymbol{\delta}(A\boldsymbol{x}_t + B\rho v_t) + \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} \\ &= \boldsymbol{\varphi}v_t + \boldsymbol{\omega}_f\boldsymbol{\delta}[A(A\boldsymbol{x}_{t-1} + Bv_t) + B\rho v_t] + \boldsymbol{\omega}_b\boldsymbol{x}_{t-1} \\ &= [\boldsymbol{\varphi} + \boldsymbol{\omega}_f\boldsymbol{\delta}(A + \rho)B]\, v_t + [\boldsymbol{\omega}_f\boldsymbol{\delta}AA + \boldsymbol{\omega}_b]\, \boldsymbol{x}_{t-1} \end{aligned}$$

They are thus equivalent if

$$\begin{aligned} B - \boldsymbol{\varphi} &= \boldsymbol{\omega}_f\boldsymbol{\delta}(AB + \rho B) \\ \boldsymbol{\omega}_b &= (I_3 - \boldsymbol{\omega}_f\boldsymbol{\delta}A)A \end{aligned} \tag{A.22}$$

for certain matrices $\boldsymbol{\omega}_b$ and $\boldsymbol{\omega}_f$

$$\boldsymbol{\omega}_b = \begin{bmatrix} \omega_{b,11} & \omega_{b,12} & \omega_{b,13} \\ \omega_{b,21} & \omega_{b,22} & \omega_{b,23} \\ \omega_{b,31} & \omega_{b,32} & \omega_{b,33} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\omega}_f = \begin{bmatrix} \omega_{f,11} & \omega_{f,12} & \omega_{f,13} \\ \omega_{f,21} & \omega_{f,22} & \omega_{f,23} \\ \omega_{f,31} & \omega_{f,32} & \omega_{f,33} \end{bmatrix}$$

The system of restrictions (A.22) implies that $\omega_{b,11} = \omega_{b,21} = \omega_{b,31} = 0$. I need to multiply the system by $\tilde{A}$ to back out the structural dynamics. In particular, we can write inflation dynamics as

$$\begin{aligned} \pi_t &= \omega_1 \pi_{t-1} + \omega_2 p_{t-1} + \kappa \tilde{y}_t + \omega_3 \mathbb{E}_t \tilde{y}_{t+1} + \omega_4 \mathbb{E}_t \pi_{t+1} + \omega_5 \mathbb{E}_t p_{t+1} \\ &= \frac{\omega_1}{1-\omega_5}\pi_{t-1} + \frac{\omega_2+\omega_5}{1-\omega_5} p_{t-1} + \frac{\kappa}{1-\omega_5}\tilde{y}_t + \frac{\omega_3}{1-\omega_5}\mathbb{E}_t \tilde{y}_{t+1} + \frac{\omega_4+\omega_5}{1-\omega_5}\mathbb{E}_t \pi_{t+1} \end{aligned} \tag{A.23}$$

□

Proof of Corollary 1. We can write (4.4) as

$$\begin{aligned} \pi_t = \omega_\pi \pi_{t-1} + \kappa \tilde{y}_t + \delta_\pi \beta \mathbb{E}_t \pi_{t+1} \\ + \omega_p p_{t-1} + (\gamma_y - 1)\kappa \tilde{y}_t + \delta_y \mathbb{E}_t \tilde{y}_{t+1} \end{aligned}$$

Using the model dynamics (3.12)-(3.13), we can write the second part as

$$\begin{aligned} \omega_p p_{t-1} + (\gamma_y - 1)\kappa \tilde{y}_t + \delta_y \mathbb{E}_t \tilde{y}_{t+1} = [\omega_p + (\gamma_y - 1)\kappa a_y + \delta_y a_y(\vartheta_1 + \vartheta_2) + \delta_y b_y] p_{t-1} \\ + [(\gamma_y - 1)\kappa b_y - \delta_y a_y \vartheta_1 \vartheta_2] p_{t-2} + [(\gamma_y - 1)\kappa c_y + \delta_y a_y \chi_\pi + \delta_y \rho \chi_y] v_t \end{aligned}$$

and we can use the two degrees of freedom to set

$$\begin{aligned} \omega_p + (\gamma_y - 1)\kappa a_y + \delta_y a_y(\vartheta_1 + \vartheta_2) + \delta_y b_y = 0 \\ (\gamma_y - 1)\kappa b_y - \delta_y a_y \vartheta_1 \vartheta_2 = 0 \end{aligned}$$

where $(\gamma_y - 1)\kappa c_y + \delta_y a_y \chi_\pi + \delta_y \rho \chi_y = \chi$. □

Proof of Proposition 5. Recall the policy functions

$$c_{it} = \frac{\beta\phi_\pi}{\sigma}\mathbb{E}_{it} p_{t-1} + \left(1 - \beta - \frac{\beta\phi_y}{\sigma}\right)\mathbb{E}_{it}\tilde{y}_t - \frac{\beta(1+\phi_\pi)}{\sigma}\mathbb{E}_{it} p_t + \frac{\beta}{\sigma}\mathbb{E}_{it} p_{t+1} - \frac{\beta}{\sigma}\mathbb{E}_{it} v_t + \beta \mathbb{E}_{it} c_{i,t+1} \tag{A.24}$$

$$p^*_{jt} = (1-\beta\theta)\mathbb{E}_{jt}p_t + \frac{\kappa\theta}{1-\theta}\mathbb{E}_{jt}\tilde{y}_t + \beta\theta\mathbb{E}_{jt}p^*_{j,t+1} \tag{A.25}$$

We now turn to solving the expectation terms. We can write the fundamental representation of the signal process as a system containing (3.8) and (3.10), which admits the following state-space representation

$$\begin{aligned} \boldsymbol{Z}_t &= \boldsymbol{F}\boldsymbol{Z}_{t-1} + \boldsymbol{\Phi}\boldsymbol{s}_{it} \\ \boldsymbol{X}_{it} &= \boldsymbol{H}\boldsymbol{Z}_t + \boldsymbol{\Psi}\boldsymbol{s}_{it} \end{aligned} \tag{A.26}$$

with $\boldsymbol{F} = \rho$, $\boldsymbol{\Phi} = \begin{bmatrix}\sigma_\varepsilon & 0\end{bmatrix}$, $\boldsymbol{Z}_t = v_t$, $\boldsymbol{s}_{lgt} = \begin{bmatrix}\varepsilon^v_t \\ u_{lgt}\end{bmatrix}$, $\boldsymbol{H} = 1$, $\boldsymbol{\Psi}_g = \begin{bmatrix}0 & \sigma_{gu}\end{bmatrix}$ and $\boldsymbol{X}_{lgt} = x_{lgt}$. It is convenient to rewrite the uncertainty parameters in terms of precision: define $\tau_\varepsilon \equiv \frac{1}{\sigma^2_\varepsilon}$ and $\tau_g \equiv \frac{1}{\sigma^2_{gu}}$. The signal system can be written as

$$\boldsymbol{X}_{igt} = \frac{\sigma_\varepsilon}{1-\rho L}\varepsilon^v_t + \sigma_{gu}u_{it} = \begin{bmatrix}\frac{\tau_\varepsilon^{-1/2}}{1-\rho L} & \tau_g^{-1/2}\end{bmatrix}\begin{bmatrix}\varepsilon^v_t \\ u_{lgt}\end{bmatrix} = \boldsymbol{M}_g(L)\boldsymbol{s}_{lgt}, \qquad \boldsymbol{s}_{lgt} \sim \mathcal{N}(0, I) \tag{A.27}$$

The Wold theorem states that there exists another representation of the signal process (A.27),

$$\boldsymbol{X}_{lgt} = \boldsymbol{B}_g(L)\boldsymbol{w}_{lgt}$$

such that $\boldsymbol{B}_g(z)$ is invertible and $\boldsymbol{w}_{lgt} \sim (0, \boldsymbol{V}_g)$ is white noise. Hence, we can write the following equivalence

$$\boldsymbol{X}_{lgt} = \boldsymbol{M}_g(L)\boldsymbol{s}_{lgt} = \boldsymbol{B}_g(L)\boldsymbol{w}_{lgt} \tag{A.28}$$

In the Wold representation of $\boldsymbol{X}_{lgt}$, observing $\{\boldsymbol{X}_{lgt}\}$ is equivalent to observing $\{\boldsymbol{w}_{lgt}\}$, and $\{\boldsymbol{X}^t_{lg}\}$ and $\{\boldsymbol{w}^t_{lg}\}$ contain the same information. Furthermore, note that the Wold representation has the property that, using the equivalence (A.5), both processes share the autocovariance generating function

$$\rho^g_{xx}(z) = \boldsymbol{M}_g(z)\boldsymbol{M}'_g(z^{-1}) = \boldsymbol{B}_g(z)\boldsymbol{V}_g\boldsymbol{B}'_g(z^{-1})$$

Given the state-space representation of the signal process (A.26), optimal expectations of the exogenous fundamental take the form of a Kalman filter

$$\mathbb{E}_{lgt}v_t = (I - \boldsymbol{K}_g\boldsymbol{H})\boldsymbol{F}\mathbb{E}_{it-1}v_{t-1} + \boldsymbol{K}_g x_{lgt} = \lambda_g\mathbb{E}_{it-1}v_{t-1} + \boldsymbol{K}_g x_{lgt}$$

where $\boldsymbol{K}_g$ is given by

$$\boldsymbol{K}_g = \boldsymbol{P}_g \boldsymbol{H}' \boldsymbol{V}_g^{-1} \tag{A.29}$$

$$\boldsymbol{P}_g = \boldsymbol{F}[\boldsymbol{P}_g - \boldsymbol{P}_g \boldsymbol{H}' \boldsymbol{V}_g^{-1} \boldsymbol{H} \boldsymbol{P}_g]\boldsymbol{F} + \boldsymbol{\Phi}\boldsymbol{\Phi}' \tag{A.30}$$

We still need to find the unknowns $\boldsymbol{B}_g(z)$ and $\boldsymbol{V}_g$. Propositions 13.1-13.4 in Hamilton (1994) provide us with these objects. Unknowns $\boldsymbol{B}_g(z)$ and $\boldsymbol{V}_g$ satisfy

$$\boldsymbol{B}_g(z) = I + \boldsymbol{H}(I - \boldsymbol{F}z)^{-1}\boldsymbol{F}\boldsymbol{K}_g$$

$$\boldsymbol{V}_g = \boldsymbol{H}\boldsymbol{P}_g\boldsymbol{H}' + \boldsymbol{\Psi}_g\boldsymbol{\Psi}_g'$$

I can write (A.30) as

$$\boldsymbol{P}_g^2 + \boldsymbol{P}_g[(1-\rho^2)\sigma_{gu}^2 - \sigma_\varepsilon^2] - \sigma_\varepsilon^2\sigma_{gu}^2 = 0 \tag{A.31}$$

from which we can infer that $\boldsymbol{P}_g$ is a scalar. Denote $k_g = \boldsymbol{P}_g^{-1}$ and rewrite (A.31) as

$$\sigma_{gu}^2\sigma_\varepsilon^2 k_g^2 = [(1-\rho^2)\sigma_{gu}^2 - \sigma_\varepsilon^2]k_g + 1 \implies k_g = \frac{\tau_\varepsilon}{2}\left\{1 - \rho^2 - \frac{\tau_g}{\tau_\varepsilon} \pm \sqrt{\left[\frac{\tau_g}{\tau_\varepsilon} - (1-\rho^2)\right]^2 + 4\frac{\tau_g}{\tau_\varepsilon}}\right\}$$

I also need to find $\boldsymbol{K}_g$. Now that we have found $\boldsymbol{P}_g$ in terms of model primitives, we can obtain $\boldsymbol{K}_g$ using condition (A.29)

$$\boldsymbol{K}_g = \frac{1}{1 + k_g\sigma_{gu}^2}$$

We can finally write λ_g as

$$\lambda_g = (I - \boldsymbol{K}_g\boldsymbol{H})\boldsymbol{F} = \frac{k_g\sigma_{gu}^2\rho}{1 + k_g\sigma_{gu}^2} = \frac{1}{2}\left[\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\varepsilon} \pm \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\varepsilon}\right)^2 - 4}\right] \tag{A.32}$$

One can show that one of the roots $\lambda_{g,[1,2]}$ lies inside the unit circle and the other lies outside as long as $\rho \in (0,1)$, which guarantees that the Kalman expectation process is stationary and unique. We set λ_g to the root that lies inside the unit circle (the one with the '$-$' sign).

Notice that I can also write $\boldsymbol{V}_g$ in terms of λ_g

$$\boldsymbol{V}_g = k^{-1} + \sigma_{gu}^2 = \frac{\rho}{\lambda_g \tau_g}$$

where I have used the identity $k_g = {}^{\lambda_g \tau_g}/_{(\rho - \lambda_g)}$. Finally, I can obtain $\boldsymbol{B}_g(z)$

$$\boldsymbol{B}_g(z) = 1 + \frac{\rho z}{(1-\rho z)(1 + k\sigma_{gu}^2)} = \frac{1 - \lambda_g z}{1 - \rho z}$$

and therefore one can verify that

$$\begin{aligned} \boldsymbol{B}_g(z)\boldsymbol{V}_g\boldsymbol{B}_g'(z^{-1}) &= \boldsymbol{M}_g(z)\boldsymbol{M}_g'(z^{-1}) \\ \frac{\rho}{\lambda_g \tau_g}\frac{(1-\lambda_g z)(z-\lambda_g)}{(1-\rho z)(z-\rho)} &= \frac{\tau_\varepsilon z}{(1-\rho z)(z-\rho)} + \tau_g \end{aligned}$$

Let us now move to the forecast of *endogenous* variables. Consider a variable $f_t = A(L)\boldsymbol{s}_{it}$. Applying the Wiener-Hopf prediction filter, we can obtain the forecast as

$$\mathbb{E}_{it} f_t = \left[A(z)\boldsymbol{M}'(z^{-1})\boldsymbol{B}(z^{-1})^{-1}\right]_+ \boldsymbol{V}^{-1}\boldsymbol{B}(z)^{-1} x_{it}$$

where $[\cdot]_+$ denotes the annihilator operator.[39]

Recall from conditions (A.24)-(A.25) that we are interested in obtaining $\mathbb{E}_{lgt} v_t$, $\mathbb{E}_{lgt} p_{t-k}$ and $\mathbb{E}_{lgt} \tilde{y}_{t-k}$, $k = \{-1, 0, 1\}$. Just as we did in the example above, we need to find the $A(z)$ polynomial for each of the forecasted variables. Let us start from the exogenous fundamental v_t to verify that the Kalman and Wiener-Hopf filters result in the same forecast. I can write the fundamental as

$$v_t = \begin{bmatrix} \frac{\tau_\varepsilon^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \boldsymbol{s}_{it} = A_v(L)\boldsymbol{s}_{it}$$

Let us now move to the endogenous variables. Let us start from the household side. We need to guess (and verify) that each firm j's policy function takes the following form[40]

$$c_{it} = h_1(L) x_{l1t}$$

39. See Online Appendix H for more details on the Wiener-Hopf prediction filter and the annihilator operator.

40. In this framework agents only observe signals. As a result, the policy function can only depend on current and past signals.

Aggregate output can then be expressed as

$$\tilde{y}_t = \int h_1(L) x_{l1t}\, dj = h_1(L) \frac{\tau_\varepsilon^{-1/2}}{1-\rho L} \varepsilon_t^v$$

Using the guesses, I have

$$\begin{aligned}
\tilde{y}_{t-k} &= \begin{bmatrix} h_1(L) L^k \frac{\tau_\varepsilon^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \boldsymbol{s}_{l1t} = A_{yk}(L) \boldsymbol{s}_{l1t} \\
c_{i,t+1}^* &= \frac{h_1(L)}{L} \boldsymbol{M}_1(L) \boldsymbol{s}_{l1t} = \begin{bmatrix} h_1(L) \frac{\tau_\varepsilon^{-1/2}}{L(1-\rho L)} & \tau_1^{-1/2} \frac{h_1(L)}{L} \end{bmatrix} \boldsymbol{s}_{l1t} = A_{i1}(L) \boldsymbol{s}_{l1t}
\end{aligned}$$

Let us now move to firms. In this case we need to guess (and verify) that each firm j's policy function takes the following form

$$p_{jt}^* = h_2(L) x_{l2t}$$

Aggregate price level can then be expressed as

$$p_t = (1-\theta) \int h_2(L) x_{l2t}\, dj + \theta p_{t-1} = (1-\theta) h_2(L) \frac{\tau_\varepsilon^{-1/2}}{(1-\rho L)(1-\theta L)} \varepsilon_t^v$$

Using the guesses, I have

$$\begin{aligned}
p_{t-k} &= \begin{bmatrix} (1-\theta) \tau_\varepsilon^{-1/2} \frac{h_2(L) L^k}{(1-\rho L)(1-\theta L)} & 0 \end{bmatrix} \boldsymbol{s}_{l2t} = A_{pk}(L) \boldsymbol{s}_{l2t} \\
p_{j,t+1}^* &= \frac{h_2(L)}{L} \boldsymbol{M}_2(L) \boldsymbol{s}_{l2t} = \begin{bmatrix} \tau_\varepsilon^{-1/2} \frac{h_2(L)}{L(1-\rho L)} & \tau_2^{-1/2} \frac{h_2(L)}{L} \end{bmatrix} \boldsymbol{s}_{l2t} = A_{i2}(L) \boldsymbol{s}_{l2t}
\end{aligned}$$

We are now armed with the necessary objects in order to obtain the three different forecasts,

$$\begin{aligned}
\mathbb{E}_{lgt} v_t &= \left[A_v(z) \boldsymbol{M}_g'(z^{-1}) \boldsymbol{B}_g(z^{-1})^{-1} \right]_+ \boldsymbol{V}_g^{-1} \boldsymbol{B}_g(z)^{-1} x_{lgt} \\
&= \left[\begin{bmatrix} \frac{\tau_\varepsilon^{-1/2}}{1-\rho z} & 0 \end{bmatrix} \begin{bmatrix} \frac{z \tau_\varepsilon^{-1/2}}{z-\rho} \\ \tau_g^{-1/2} \end{bmatrix} \frac{z-\rho}{z-\lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{it} \\
&= \left[\frac{z}{(1-\rho z)(z-\lambda_g)} \right]_+ \frac{\lambda \tau_g}{\rho \tau_\varepsilon} \frac{1-\rho z}{1-\lambda_g z} x_{it} \\
&= \left[\frac{\phi_v(z)}{z-\lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon} \frac{1-\rho z}{1-\lambda_g z} x_{it}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\phi_v(z) - \phi_v(\lambda_g)}{z - \lambda_g} \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon} \frac{1 - \rho z}{1 - \lambda_g z} x_{it} \\
&= \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon (1 - \rho \lambda_g)} \frac{1}{1 - \lambda_g z} x_{it} = G_{1g}(z) x_{it}
\end{aligned} \tag{A.33}
$$

$$
\begin{aligned}
\mathbb{E}_{lgt} \tilde{y}_{t-k} &= \left[A_{yk}(z) \boldsymbol{M}_g'(z^{-1}) \boldsymbol{B}_g(z^{-1})^{-1} \right]_+ \boldsymbol{V}_g^{-1} \boldsymbol{B}_g(z)^{-1} x_{lgt} \\
&= \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} \frac{h_1(z) z^k}{1 - \rho z} & 0 \end{bmatrix} \begin{bmatrix} \frac{z \tau_\varepsilon^{-1/2}}{z - \rho} \\ \tau_g^{-1/2} \end{bmatrix} \frac{z - \rho}{z - \lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \left[\frac{h_1(z) z^{k+1}}{(1 - \rho z)(z - \lambda_g)} \right]_+ \frac{\lambda_g \tau_g}{\tau_\varepsilon \rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \left[\frac{\phi_y(z)}{z - \lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\tau_\varepsilon \rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \frac{\phi_y(z) - \phi_y(\lambda_g)}{z - \lambda_g} \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon} \left[h_1(z) z^{k+1} - h_1(\lambda_g) \lambda_g^{k+1} \frac{1 - \rho z}{1 - \rho \lambda_g} \right] \frac{1}{(1 - \lambda_g z)(z - \lambda_g)} x_{lgt} = G_{2gk}(z) x_{lgt}
\end{aligned} \tag{A.34}
$$

$$
\begin{aligned}
\mathbb{E}_{lgt} p_{t-k} &= \left[A_{pk}(z) \boldsymbol{M}_g'(z^{-1}) \boldsymbol{B}_g(z^{-1})^{-1} \right]_+ \boldsymbol{V}_g^{-1} \boldsymbol{B}_g(z)^{-1} x_{lgt} \\
&= \left[\begin{bmatrix} (1 - \theta) \tau_\varepsilon^{-1/2} \frac{h_2(z) z^k}{(1 - \rho z)(1 - \theta z)} & 0 \end{bmatrix} \begin{bmatrix} \frac{z \tau_\varepsilon^{-1/2}}{z - \rho} \\ \tau_g^{-1/2} \end{bmatrix} \frac{z - \rho}{z - \lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \left[\frac{h_2(z) z^{k+1}}{(1 - \rho z)(z - \lambda_g)(1 - \theta z)} \right]_+ \frac{(1 - \theta) \lambda_g \tau_g}{\tau_\varepsilon \rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \left[\frac{\phi_\pi(z)}{z - \lambda_g} \right]_+ \frac{(1 - \theta) \lambda_g \tau_g}{\tau_\varepsilon \rho} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= \frac{\phi_\pi(z) - \phi_\pi(\lambda_g)}{z - \lambda_g} \frac{(1 - \theta) \lambda_g \tau_g}{\rho \tau_\varepsilon} \frac{1 - \rho z}{1 - \lambda_g z} x_{lgt} \\
&= (1 - \theta) \frac{\lambda_g \tau_g}{\rho \tau_\varepsilon} \left[\frac{h_2(z) z^{k+1}}{1 - \theta z} - h_2(\lambda_g) \lambda_g^{k+1} \frac{1 - \rho z}{(1 - \rho \lambda_g)(1 - \theta \lambda_g)} \right] \frac{1}{(1 - \lambda_g z)(z - \lambda_g)} x_{lgt} = G_{3gk}(z) x_{lgt}
\end{aligned} \tag{A.35}
$$

$$
\mathbb{E}_{lgt} a_{lg,t+1} = \left[A_{ig}(z) \boldsymbol{M}_g'(z^{-1}) \boldsymbol{B}_g(z^{-1})^{-1} \right]_+ \boldsymbol{V}_g^{-1} \boldsymbol{B}_g(z)^{-1} x_{lgt}
$$

$$
\begin{aligned}
&= \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} \frac{h_g(z)}{z(1-\rho z)} & \tau_g^{-1/2} \frac{h_g(z)}{z} \end{bmatrix} \begin{bmatrix} \frac{z\tau_\varepsilon^{-1/2}}{z-\rho} \\ \tau_g^{-1/2} \end{bmatrix} \frac{z-\rho}{z-\lambda_g} \right]_+ \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= \left[\frac{h_g(z)}{\tau_\varepsilon (1-\rho z)(z-\lambda_g)} + \frac{h_g(z)(z-\rho)}{\tau_g z (z-\lambda_g)} \right]_+ \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= \left\{ \left[\frac{h_g(z)}{\tau_\varepsilon (1-\rho z)(z-\lambda_g)} \right]_+ + \left[\frac{h_g(z)(z-\rho)}{\tau_g z (z-\lambda_g)} \right]_+ \right\} \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= \left\{ \left[\frac{\phi_{ig,1}(z)}{z-\lambda_g} \right]_+ + \left[\frac{\phi_{ig,2}(z)}{z(z-\lambda_g)} \right]_+ \right\} \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= \left\{ \frac{\phi_{ig,1}(z) - \phi_{ig,1}(\lambda_g)}{z-\lambda_g} + \frac{\phi_{ig,2}(z) - \phi_{ig,2}(\lambda_g)}{\lambda_g (z-\lambda_g)} - \frac{\phi_{ig,2}(z) - \phi_{ig,2}(0)}{\lambda_g z} \right\} \frac{\lambda_g \tau_g}{\rho} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= \frac{\lambda_g}{\rho} \left\{ \frac{h_g(z)}{z-\lambda_g} \left[\frac{\tau_g}{\tau_\varepsilon (1-\rho z)} + \frac{z-\rho}{z} \right] - \frac{h_g(\lambda_g)}{z-\lambda_g} \left[\frac{\tau_g}{\tau_\varepsilon (1-\rho \lambda_g)} + \frac{\lambda_g - \rho}{\lambda_g} \right] - \frac{\rho h_g(0)}{\lambda_g z} \right\} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} \\
&= G_{4g}(z) x_{lgt} \qquad \text{(A.36)}
\end{aligned}
$$

where $\mathbb{E}_{l1t} a_{l1,t+1} = \mathbb{E}_{it} c_{i,t+1}$, $\mathbb{E}_{l2t} a_{l2,t+1} = \mathbb{E}_{jt} p^*_{j,t+1}$, and

$$
\phi_v(z) = \frac{z}{1-\rho z}, \quad \phi_\pi(z) = \frac{h_2(z) z^{k+1}}{(1-\rho z)(1-\theta z)}, \quad \phi_y(z) = \frac{h_1(z) z^{k+1}}{1-\rho z}
$$
$$
\phi_{ig,1}(z) = \frac{h_g(z)}{\tau_\varepsilon (1-\rho z)}, \quad \phi_{ig,2}(z) = \frac{h_g(z)(z-\rho)}{\tau_g}
$$

Rearranging terms, we obtain (A.37)-(A.41). We can show that expectations satisfy

$$
\mathbb{E}_{lgt} v_t = \left(1 - \frac{\lambda_g}{\rho}\right) \frac{1}{1-\lambda_g z} x_{lgt} = G_{1g}(z) x_{lgt} \qquad \text{(A.37)}
$$

$$
\mathbb{E}_{lgt} a_{k,t-1} = (1-\theta_k) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_k(z) z^2 (1-\rho\lambda_g)}{1-\theta_k z} - \frac{h_k(\lambda_g)\lambda_g^2 (1-\rho z)}{1-\theta_k \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)} x_{lgt} = G_{2k}(z) x_{lgt} \qquad \text{(A.38)}
$$

$$
\mathbb{E}_{lgt} a_{k,t} = (1-\theta_k) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_k(z) z (1-\rho\lambda_g)}{1-\theta_k z} - \frac{h_k(\lambda_g)\lambda_g (1-\rho z)}{1-\theta_k \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)} x_{lgt} = G_{3k}(z) x_{lgt} \qquad \text{(A.39)}
$$

$$
\mathbb{E}_{lgt} a_{k,t+1} = (1-\theta_k) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_k(z)(1-\rho\lambda_g)}{1-\theta_k z} - \frac{h_k(\lambda_g)(1-\rho z)}{1-\theta_k \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)} x_{lgt} = G_{4k}(z) x_{lgt} \qquad \text{(A.40)}
$$

$$
\mathbb{E}_{lgt} a_{lg,t+1} = \left\{ \frac{h_g(z)}{z-\lambda_g} \left[\left(1 - \frac{\lambda_g}{\rho}\right) \frac{1-\rho\lambda_g}{1-\rho z} + \frac{\lambda_g (z-\rho)}{\rho z} \right] - \frac{h_g(0)}{z} \right\} \frac{1-\rho z}{1-\lambda_g z} x_{lgt} = G_{5g}(z) x_{lgt} \qquad \text{(A.41)}
$$

Recall the best response for household i and firm j, conditions (A.24)-(A.25). In order to be consistent with agent optimization, the policy functions $h_g(z)$ must satisfy (A.24)-(A.25) at all times and signals. Plugging the obtained expressions, we can write

$$
\begin{aligned}
a_{lgt} &= \varphi_g \mathbb{E}_{lgt} v_t + \beta_g \mathbb{E}_{lgt} a_{lg,t+1} + \sum_{j=1}^{2} \mu_{gj} \mathbb{E}_{lgt} a_{j,t-1} + \sum_{j=1}^{2} \gamma_{gj} \mathbb{E}_{lgt} a_{j,t} + \sum_{j=1}^{2} \alpha_{gj} \mathbb{E}_{lgt} a_{j,t+1} \\
h_g(z) x_{lgt} &= \varphi_g G_{1g}(z) x_{lgt} + \beta_g G_{5g}(z) x_{lgt} + \sum_{j=1}^{2} \mu_{gj} G_{2j}(z) x_{lgt} + \sum_{j=1}^{2} \gamma_{gj} G_{3j}(z) x_{lgt} + \sum_{j=1}^{2} \alpha_{gj} G_{4j}(z) x_{lgt} \\
h_g(z) &= \varphi_g G_{1g}(z) + \beta_g G_{5g}(z) + \sum_{j=1}^{2} \mu_{gj} G_{2j}(z) + \sum_{j=1}^{2} \gamma_{gj} G_{3j}(z) + \sum_{j=1}^{2} \alpha_{gj} G_{4j}(z) \\
&= \varphi_g \left(1 - \frac{\lambda_g}{\rho}\right) \frac{1}{1-\lambda_g z} + \beta_g \left\{ \frac{h_g(z)}{z-\lambda_g} \left[\left(1 - \frac{\lambda_g}{\rho}\right) \frac{1-\rho\lambda_g}{1-\rho z} + \frac{\lambda_g(z-\rho)}{\rho z} \right] - \frac{h_g(0)}{z} \right\} \frac{1-\rho z}{1-\lambda_g z} \\
&+ \sum_{j=1}^{2} \mu_{gj}(1-\theta_j) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_j(z) z^2 (1-\rho\lambda_g)}{1-\theta_j z} - \frac{h_j(\lambda_g)\lambda_g^2(1-\rho z)}{1-\theta_j \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)} \\
&+ \sum_{j=1}^{2} \gamma_{gj}(1-\theta_j) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_j(z) z (1-\rho\lambda_g)}{1-\theta_j z} - \frac{h_j(\lambda_g)\lambda_g(1-\rho z)}{1-\theta_j \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)} \\
&+ \sum_{j=1}^{2} \alpha_{gj}(1-\theta_j) \left(1 - \frac{\lambda_g}{\rho}\right) \left[\frac{h_j(z)(1-\rho\lambda_g)}{1-\theta_j z} - \frac{h_j(\lambda_g)(1-\rho z)}{1-\theta_j \lambda_g} \right] \frac{1}{(1-\lambda_g z)(z-\lambda_g)}
\end{aligned}
$$

where

$$
\begin{aligned}
\varphi_1 &= -\frac{\beta}{\sigma} & \varphi_2 &= 0 \\
\beta_1 &= \beta & \beta_2 &= \beta\theta \\
\mu_{11} &= 0 & \mu_{21} &= 0 \\
\mu_{12} &= \frac{\beta\phi_\pi}{\sigma} & \mu_{22} &= 0 \\
\gamma_{11} &= 1 - \beta\left(1 + \frac{\phi_y}{\sigma}\right) & \gamma_{21} &= \frac{\kappa\theta}{1-\theta} \\
\gamma_{12} &= -\frac{\beta(1+\phi_\pi)}{\sigma} & \gamma_{22} &= 1 - \beta\theta \\
\alpha_{11} &= 0 & \alpha_{21} &= 0 \\
\alpha_{12} &= \frac{\beta}{\sigma} & \alpha_{22} &= 0 \\
\theta_1 &= 0 & \theta_2 &= \theta
\end{aligned}
$$

Multiplying both sides by $z(z-\lambda_g)(1-\lambda_g z)(1-\theta_1 z)(1-\theta_2 z)$ we obtain

$$
\begin{aligned}
&h_g(z)z(z-\lambda_g)(1-\lambda_g z)(1-\theta_1 z)(1-\theta_2 z) = \\
&= \varphi_g\left(1-\frac{\lambda_g}{\rho}\right) z(z-\lambda_g)(1-\theta_1 z)(1-\theta_2 z) \\
&+ \beta_g\left\{h_g(z)\left[\left(1-\frac{\lambda_g}{\rho}\right)(1-\rho\lambda_g)z + \frac{\lambda_g}{\rho z}(z-\rho)(1-\rho z)\right] - h_g(0)(z-\lambda_g)(1-\rho z)\right\}(1-\theta_1 z)(1-\theta_2 z) \\
&+ \sum_{j=1}^{2}\mu_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)\left[h_j(z)z^3(1-\rho\lambda_g)(1-\theta_{\neg j}z) - \frac{h_j(\lambda_g)\lambda_g^2 z(1-\rho z)(1-\theta_1 z)(1-\theta_2 z)}{1-\theta_j\lambda_g}\right] \\
&+ \sum_{j=1}^{2}\gamma_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)\left[h_j(z)z^2(1-\rho\lambda_g)(1-\theta_{\neg j}z) - \frac{h_j(\lambda_g)\lambda_g z(1-\rho z)(1-\theta_1 z)(1-\theta_2 z)}{1-\theta_j\lambda_g}\right] \\
&+ \sum_{j=1}^{2}\alpha_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)\left[h_j(z)z(1-\rho\lambda_g)(1-\theta_{\neg j}z) - \frac{h_j(\lambda_g)z(1-\rho z)(1-\theta_1 z)(1-\theta_2 z)}{1-\theta_j\lambda_g}\right]
\end{aligned}
$$

Rearranging the LHS by $h_g(z)$,

$$
\begin{aligned}
h_g(z)\Big\{&z(z-\lambda_g)(1-\lambda_g z)(1-\theta_1 z)(1-\theta_2 z) \\
&- \beta_g\left[\left(1-\frac{\lambda_g}{\rho}\right)(1-\rho\lambda_g)z + \frac{\lambda_g}{\rho z}(z-\rho)(1-\rho z)\right](1-\theta_1 z)(1-\theta_2 z)\Big\} \\
&- \sum_{j=1}^{2}\mu_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)z^3(1-\rho\lambda_g)(1-\theta_{\neg j}z)h_j(z) \\
&- \sum_{j=1}^{2}\gamma_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)z^2(1-\rho\lambda_g)(1-\theta_{\neg j}z)h_j(z) \\
&- \sum_{j=1}^{2}\alpha_{gj}(1-\theta_j)\left(1-\frac{\lambda_g}{\rho}\right)z(1-\rho\lambda_g)(1-\theta_{\neg j}z)h_j(z)
\end{aligned}
$$

and the RHS can be rewritten as

$$
\begin{aligned}
d_g(z) = \varphi_g&\left(1-\frac{\lambda_g}{\rho}\right)z(z-\lambda_g)(1-\theta_1 z)(1-\theta_2 z) - h_g(0)\beta_g(z-\lambda_g)(1-\rho z)(1-\theta_1 z)(1-\theta_2 z) \\
&- \left\{\left(1-\frac{\lambda_g}{\rho}\right)\sum_{j=1}^{2}\frac{1-\theta_j}{1-\theta_j\lambda_g}[\mu_{gj}\lambda_g^2 + \gamma_{gj}\lambda_g + \alpha_{gj}]h_j(\lambda_g)\right\}z(1-\rho z)(1-\theta_1 z)(1-\theta_2 z)
\end{aligned}
$$

We can write the system in matrix form as

$$\boldsymbol{C}(z)\boldsymbol{h}(z) = \boldsymbol{d}(z)$$

where

$$\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix}, \qquad \boldsymbol{h}(z) = \begin{bmatrix} h_1(z) \\ h_2(z) \end{bmatrix}, \qquad \boldsymbol{d}(z) = \begin{bmatrix} d_1(z) \\ d_2(z) \end{bmatrix}$$

$$\begin{aligned}
C_{gg}(z) &= (z-\beta_g)(z-\lambda_g)(1-\lambda_g z)(1-\theta_1 z)(1-\theta_2 z) \\
&\qquad - (1-\theta_g)\left(1-\frac{\lambda_g}{\rho}\right)(1-\rho\lambda_g)(1-\theta_{\neg g} z)z(\mu_{gg}z^2+\gamma_{gg}z+\alpha_{gg}) \\
C_{gn}(z) &= -(1-\theta_n)\left(1-\frac{\lambda_g}{\rho}\right)(1-\rho\lambda_g)(1-\theta_g z)(\mu_{gn}z^3+\gamma_{gn}z^2+\alpha_{gn}z) \\
d_g(z) &= \left[\varphi_g\left(1-\frac{\lambda_g}{\rho}\right)z(z-\lambda_g) - h_g(0)\beta_g(z-\lambda_g)(1-\rho z) - \tilde{h}_g z(1-\rho z)\right](1-\theta_1 z)(1-\theta_2 z)
\end{aligned}$$

Cancelling out parameters equal to zero to simplify the expressions, we can write

$$\begin{aligned}
C_{11}(z) &= \left[(z-\beta_1)(z-\lambda_1)(1-\lambda_1 z) - \left(1-\frac{\lambda_1}{\rho}\right)(1-\rho\lambda_1)\gamma_{11}z^2\right](1-\theta_2 z) \\
C_{12}(z) &= -(1-\theta_2)\left(1-\frac{\lambda_1}{\rho}\right)(1-\rho\lambda_1)z(\mu_{12}z^2+\gamma_{12}z+\alpha_{12}) \\
C_{21}(z) &= -\left(1-\frac{\lambda_2}{\rho}\right)(1-\rho\lambda_2)(1-\theta_2 z)\gamma_{21}z^2 \\
C_{22}(z) &= (z-\beta_2)(z-\lambda_2)(1-\lambda_2 z)(1-\theta_2 z) - (1-\theta_2)\left(1-\frac{\lambda_2}{\rho}\right)(1-\rho\lambda_2)\gamma_{22}z^2 \\
d_1(z) &= \left[\varphi_1\left(1-\frac{\lambda_1}{\rho}\right)z(z-\lambda_1) - h_1(0)\beta_1(z-\lambda_1)(1-\rho z) - \tilde{h}_1 z(1-\rho z)\right](1-\theta_2 z) \\
d_2(z) &= \left[-h_g(0)\beta_2(z-\lambda_2)(1-\rho z) - \tilde{h}_2 z(1-\rho z)\right](1-\theta_2 z)
\end{aligned}$$

and the solution to the policy functions is given by

$$\boldsymbol{h}(z) = \boldsymbol{C}(z)^{-1}\boldsymbol{d}(z) = \frac{\text{adj } \boldsymbol{C}(z)}{\det \boldsymbol{C}(z)}\boldsymbol{d}(z)$$

Note that the degree of $\boldsymbol{C}(z)$ is 8, given that $\theta_1 = 0$. Denote the inside roots of $\det \boldsymbol{C}(z)$ as $\{\zeta_1, \zeta_2, ..., \zeta_{n_1}\}$ and the outside roots as $\{\vartheta_1^{-1}, \vartheta_2^{-1}, ..., \vartheta_{n_1}^{-1}\}$. Because agents cannot use future signals, the inside roots have to be removed. Note that the number of free constants

in $\boldsymbol{d}$ is 4: $\{h_g(0), \tilde{h}_g\}_{g=1}^2$. For a unique solution, it must be the case that the number of outside roots is $n_2 = 4$. Also note that by Cramer's rule, $h_g(z)$ is given by

$$h_1(z) = \frac{\det \begin{bmatrix} d_1(z) & C_{12}(z) \\ d_2(z) & C_{22}(z) \end{bmatrix}}{\det \boldsymbol{C}(z)}, \qquad h_2(z) = \frac{\det \begin{bmatrix} C_{11}(z) & d_1(z) \\ C_{21}(z) & d_2(z) \end{bmatrix}}{\det \boldsymbol{C}(z)}$$

The degree of the numerator is 7, as the highest degree of $d_g(z)$ is 1 degree less than $C_{gg}(z)$. By choosing the constants $\{h_g(0), \tilde{h}_g\}_{g=1}^2$, the 4 inside roots will be removed. Therefore, the 4 constants are solutions to the following system of linear equations[41]

$$\det \begin{bmatrix} d_1(\zeta_n) & C_{12}(\zeta_n) \\ d_2(\zeta_n) & C_{22}(\zeta_n) \end{bmatrix} = 0, \qquad \text{for } \{\zeta_n\}_{n=1}^4$$

where $n_2 = 4$. After removing the inside roots in the denominator, the degree of the numerator is 3 and the degree of the denominator is 4. As a result, the solution to $h_g(z)$ takes the form

$$h_g(z) = \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z + \widetilde{\psi}_{g3} z^2 + \widetilde{\psi}_{g4} z^3}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)(1-\vartheta_4 z)}$$

Given the model conditions, we have that $\vartheta_4 = \theta$. We can write

$$\begin{aligned} h_g(z) &= \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z + \widetilde{\psi}_{g3} z^2 + \widetilde{\psi}_{g4} z^3}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)(1-\theta z)} \\ &= \frac{\widetilde{\psi}_{g4}(z-\eta_{g1})(z-\eta_{g2})(z-\eta_{g3})}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)(1-\theta z)} \\ &= \frac{-\widetilde{\psi}_{g4}\eta_{g1}\eta_{g2}\eta_{g3}(1-\eta_{g1}^{-1} z)(1-\eta_{g2}^{-1} z)(1-\eta_{g3}^{-1} z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)(1-\theta z)} \\ &= \frac{-\widetilde{\psi}_{g4}\eta_{g1}\eta_{g2}\eta_{g3}(1-\xi_{g1} z)(1-\xi_{g2} z)(1-\xi_{g3} z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)(1-\theta z)} \end{aligned}$$

where $(\eta_{g1}, \eta_{g2}, \eta_{g3})$ are the roots of $\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z + \widetilde{\psi}_{g3} z^2 + \widetilde{\psi}_{g4} z^3$. We also have that $\xi_{13} = \xi_{22} = \xi_{23} = \theta$. Hence, we can write

$$\tilde{y}_t = h_1(z) v_t$$

41. The set of constants that solve the system of equations for $h_1(z)$ also solves it for $h_2(z)$, since $\{\zeta_n\}_{n=1}^4$ are roots of $\det \boldsymbol{C}(z)$, leaving vectors in $\boldsymbol{C}(\zeta_n)$ being linearly dependent.

$$
\begin{aligned}
&= \frac{-\widetilde{\psi}_{14}\eta_{11}\eta_{12}\eta_{13}(1-\xi_{11}z)(1-\xi_{12}z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \phi_1 \frac{(1-\xi_{11}z)(1-\xi_{12}z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \psi_{11}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 z}v_t + \psi_{12}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 z}v_t + \psi_{13}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_3 z}v_t \\
&= \psi_{11}\widetilde{\vartheta}_{1t} + \psi_{12}\widetilde{\vartheta}_{2t} + \psi_{13}\widetilde{\vartheta}_{3t} \\
p_t &= (1-\theta)h_2(z)\frac{1}{1-\theta z}v_t \\
&= \frac{-\widetilde{\psi}_{24}\eta_{21}\eta_{22}\eta_{23}(1-\theta)(1-\xi_{21}z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \phi_2 \frac{1-\xi_{21}z}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \psi_{21}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 z}v_t + \psi_{22}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 z}v_t + \psi_{23}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_3 z}v_t \\
&= \psi_{21}\widetilde{\vartheta}_{1t} + \psi_{22}\widetilde{\vartheta}_{2t} + \psi_{23}\widetilde{\vartheta}_{3t}
\end{aligned}
$$

Using $\pi_t = (1-L)p_t$, we can write

$$
\begin{aligned}
\pi_t &= (1-\theta)h_2(z)\frac{1-z}{1-\theta z}v_t \\
&= \frac{-\widetilde{\psi}_{24}\eta_{21}\eta_{22}\eta_{23}(1-\theta)(1-\xi_{21}z)(1-z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \phi_2 \frac{(1-\xi_{21}z)(1-z)}{(1-\vartheta_1 z)(1-\vartheta_2 z)(1-\vartheta_3 z)}v_t \\
&= \psi_{31}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 z}v_t + \psi_{32}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 z}v_t + \psi_{33}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_3 z}v_t \\
&= \psi_{31}\widetilde{\vartheta}_{1t} + \psi_{32}\widetilde{\vartheta}_{2t} + \psi_{33}\widetilde{\vartheta}_{3t}
\end{aligned}
$$

We can finally write

$$
\begin{aligned}
\boldsymbol{a}_t &= \begin{bmatrix} \tilde{y}_t \\ p_t \\ \pi_t \end{bmatrix} \\
&= Q\widetilde{\vartheta}_t
\end{aligned}
$$

$$= \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \begin{bmatrix} \widetilde{\vartheta}_{1t} \\ \widetilde{\vartheta}_{2t} \\ \widetilde{\vartheta}_{3t} \end{bmatrix}$$

Notice that we can write

$$\widetilde{\vartheta}_{kt}(1 - \vartheta_k L) = \left(1 - \frac{\vartheta_k}{\rho}\right) v_t \implies \widetilde{\vartheta}_{kt} = \vartheta_k \widetilde{\vartheta}_{k,t-1} + \left(1 - \frac{\vartheta_k}{\rho}\right) v_t$$

which we can write as a system as

$$\widetilde{\vartheta}_t = \Lambda \widetilde{\vartheta}_{t-1} + \Gamma v_t$$

where

$$\Lambda = \begin{bmatrix} \vartheta_1 & 0 & 0 \\ 0 & \vartheta_2 & 0 \\ 0 & 0 & \vartheta_3 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1 - \frac{\vartheta_1}{\rho} \\ 1 - \frac{\vartheta_2}{\rho} \\ 1 - \frac{\vartheta_3}{\rho} \end{bmatrix}$$

Hence, we can write

$$\begin{aligned} \boldsymbol{a}_t &= Q\widetilde{\theta}_t \\ &= Q(\Lambda \widetilde{\theta}_{t-1} + \Gamma \xi_t) \\ &= Q\Lambda \widetilde{\theta}_{t-1} + Q\Gamma \xi_t \\ &= Q\Lambda Q^{-1} \boldsymbol{a}_{t-1} + Q\Gamma \xi_t \\ &= A\boldsymbol{a}_{t-1} + B\xi_t \end{aligned} \tag{A.42}$$

□

B Robustness on Inflation Persistence and Information Frictions

B.1 Inflation Persistence

We begin our robustness analysis by considering alternative inflation measures. Figure A.1 presents the CPI and PCE series (together with the GDP Deflator growth). All inflation measures are closely correlated. I report the correlation matrix across different sub-sample periods in Table A.i. The three main inflation measures exhibit a high and positive corre-

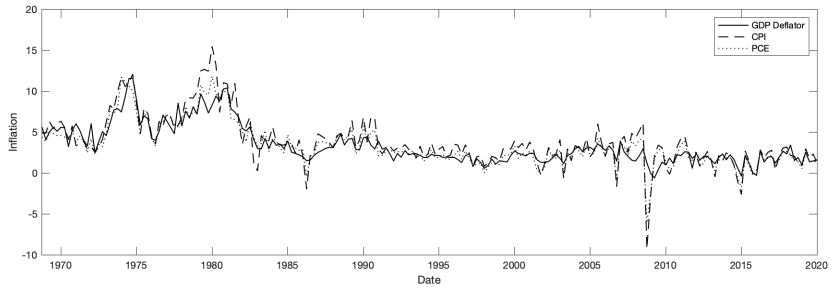


Figure A.1: Time Series of GDP Deflator, CPI and PCE.

	1969-2020		
Variable	GDP Deflator	CPI	PCE
GDP Deflator	1.00		
CPI	0.86	1.00	
PCE	0.91	0.96	1.00
	1969-1985		
GDP Deflator	1.00		
CPI	0.83	1.00	
PCE	0.88	0.92	1.00
	1985-2020		
GDP Deflator	1.00		
CPI	0.66	1.00	
PCE	0.73	0.96	1.00

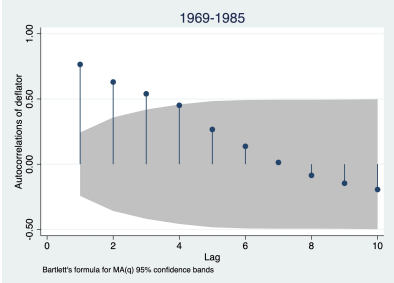
Table A.i: Correlation matrix

lation in the pre-1985 period. In the post-1985 period, there is a detachment between the GDP deflator and the two other price measures, CPI and PCE, which still exhibit a high degree of correlation.

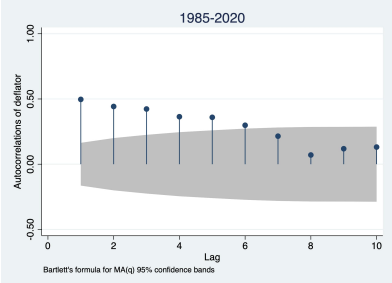
We repeat the structural break analysis discussed in the main body for CPI and PCE inflation, and we find similar results in Table A.ii, with the structural change in dynamics being less evident in the core series.

Autocorrelation Function Let us start from the most agnostic analysis of inflation persistence. Figure A.2 plots the autocorrelation function for the three main inflation measures across subsamples. Focusing on the second and third columns, I find a significant fall in the first-order autocorrelation for the three measures. For instance, the first-order autocorrelation for all inflation measures in the pre-1985 sample is around 0.75, while the same statistic for the second period ranges from 0.5 to 0.3 depending on the measure.

(a) GDP Deflator, 1947-1985

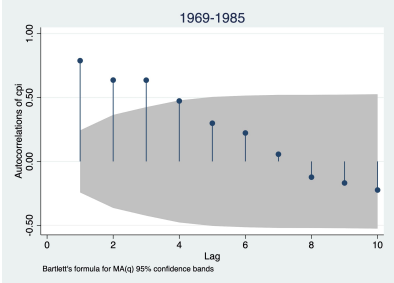


(b) GDP Deflator, 1969-1985

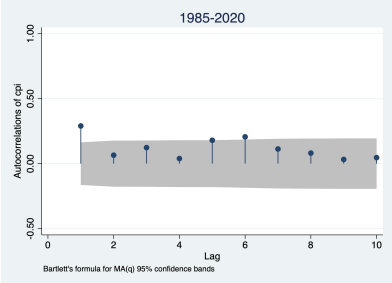


(c) GDP Deflator, 1985-2020

(d) CPI, 1947-1985

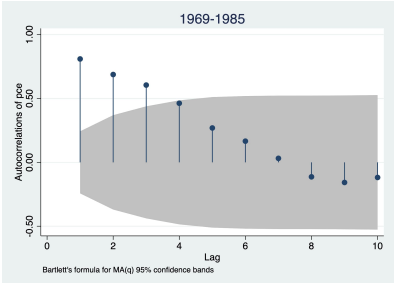


(e) CPI, 1969-1985

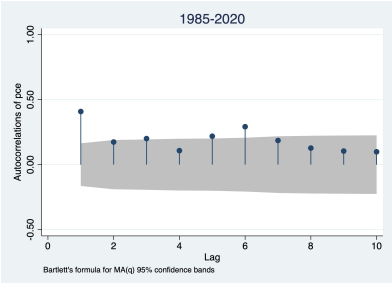


(f) CPI, 1985-2020

(g) PCE, 1947-1985



(h) PCE, 1969-1985



(i) PCE, 1985-2020

Figure A.2: Autocorrelation function of GDP Deflator (first row), CPI (second row) and PCE (last row)

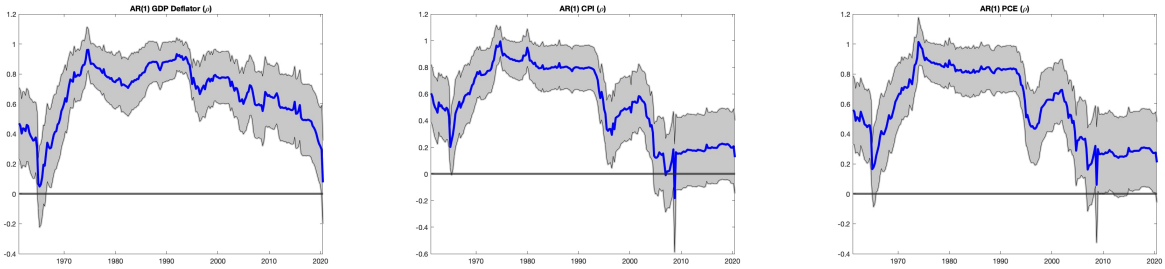
	(1) CPI	(2) PCE
π_{t-1}	0.793***	0.837***
	(0.0827)	(0.0672)
$\pi_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$	-0.497***	-0.434***
	(0.143)	(0.117)
Constant	1.396**	0.990**
	(0.542)	(0.431)
Constant$\times \mathbb{1}_{\{t \geq t^*\}}$	0.370	0.283
	(0.607)	(0.477)
Observations	206	206

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.ii: Regression table

Rolling Sample I compute rolling-sample estimates of an independent AR(1) process using a 14-year window for the different inflation measures. Figure A.3 plots the time-varying persistence parameter ρ_t with 95% confidence bands. The results suggest that there is significant time variation in inflation persistence.

Time-Varying Parameter Autorregression We assume that the persistence coefficient in the AR(1) process follows a Random Walk: $\rho_{t+1} = \rho_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u)$, where the model is estimated using Bayesian methods. The model is estimated using Bayesian methods. Our prior selection is standard, following Nakajima (2011), using the invert Wishart and



(a) GDP Deflator (b) CPI (c) PCE

Figure A.3: First-order autocorrelation of GDP Deflator, CPI and PCE, rolling sample (14y window)

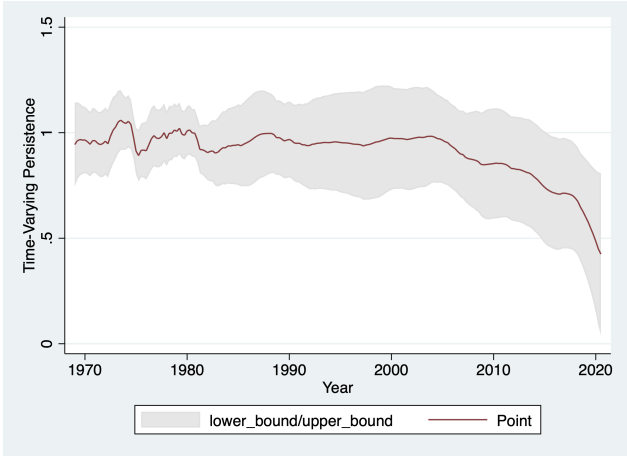


Figure A.4: Time-varying persistence.

invert Gamma distributions

$$\rho_1 \sim \mathcal{N}(0, 10 \times I), \quad \sigma_\varepsilon \sim \mathcal{IG}(2, 0.02), \quad \Sigma_u \sim \mathcal{IW}(4, 40 \times I)$$

I plot the estimated ρ_t with 95% confidence bands in Figure A.4. The fall in persistence is delayed until the mid 2000s, but the overall fall is consistent with our previous findings.

Unit Root Tests Inspecting Figure A.3, one could imagine that inflation is characterized by a unit root process in the pre-1985 sample and not afterwards. In order to investigate this, I proceed via a cross-sample unit root analysis using both the Augmented Dickie-Fuller and the Phillips-Perron tests. I report our results in Table A.iii, including the p-values of both unit root tests under the null of unit root. Focusing on the last two rows I find that, consistent with our previous evidence on the first-order autocorrelation, the null hypothesis of a unit root series cannot be rejected by any of the unit root tests conducted in the different inflation measures in the pre-1985 period. On the other hand, when I repeat the similar analysis in the post-1985 period, I find a strong rejection of the null hypothesis, suggesting that inflation can no longer be described as a unit root process. Having understood the close relation between the roots of the inflation dynamic process and its persistence, I can conclude that inflation persistence fell in the post-1985 period.

p-values, null = series has unit root		
1969-2020		
Variable	ADF	Phillips-Perron
GDP Deflator	0.23	0.02
CPI	0.11	0.00
PCE	0.16	0.00
1969-1985		
Variable	ADF	Phillips-Perron
GDP Deflator	0.15	0.07
CPI	0.17	0.09
PCE	0.055	0.09
1985-2020		
Variable	ADF	Phillips-Perron
GDP Deflator	0.07	0.00
CPI	0.00	0.00
PCE	0.01	0.00

Table A.iii: Unit Root Tests for Inflation Measures.

Dominant Root A further procedure of studying persistence that relies on the roots of the dynamic process of inflation is the dominant root analysis. Consider the AR(p) process

$$\pi_t = \rho_1 \pi_{t-1} + \rho_2 \pi_{t-2} + \ldots + \rho_p \pi_{t-p} + \varepsilon_t^{\pi}$$

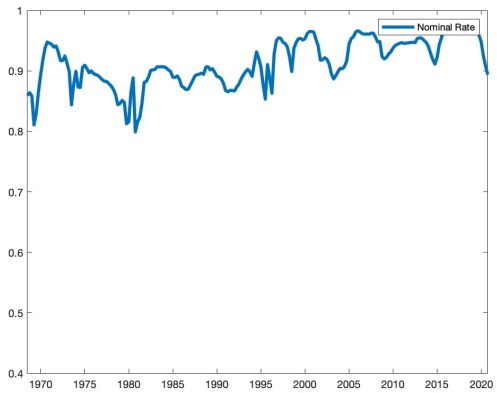
with companion matrix $R(p)$. The root of the characteristic polynomial of $R(p)$ with the largest magnitude is the dominant root of interest. Notice that in the case of an AR(p) where $p > 1$, the dominant root will depend not only on the first lag coefficient but in all of them. An AR(p) is considered to be stable if all the roots of the characteristic polynomial of matrix $R(p)$ have an absolute value lower than 1. One can therefore proceed as in the unit root case, and study the dominant root of the underlying inflation process over the different subsamples. We find that the dominant root in the 1968:Q4-1984:Q4 period is 0.870 and 0.841 in the 1985:Q1-2020:Q1 period, suggesting a moderate fall in persistence.

B.2 Monetary Policy Shock Process

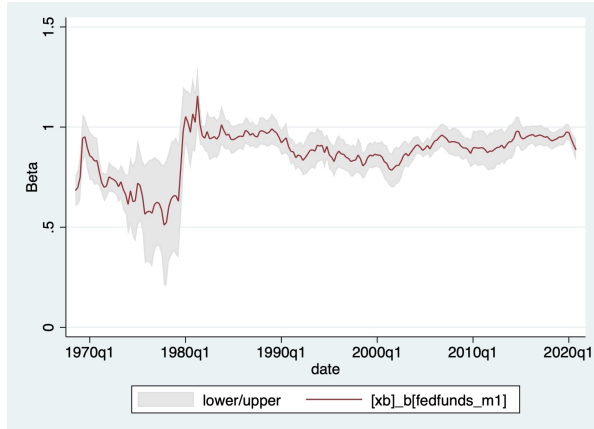
B.2.1 Persistence

Rolling Sample AR(1) Estimate of Persistence In order to test for changes in monetary shocks' persistence, I check nominal interest rates' autocorrelation. I plot in Figure A.5a the rolling (14 years) first-order autocorrelation of the nominal Fed Funds rate. I find no evidence for a fall in the first-order autocorrelation over time.

As a robustness check, I estimate (D.3) and plot the rolling estimate $\hat{\rho}_v$ over time. Notice



(a) Non-parametric First-order Autocorrelation



(b) First-order Autocorrelation, GMM

Figure A.5: First-order autocorrelation of Nominal interest rates

that the NK model implies that the error term ξ_t in (D.3) is serially correlated. In fact, the NK model suggests that ξ_t follows an ARMA(1,1), or equivalently an MA(∞),

$$\begin{aligned}\xi_t &= \rho_v \xi_{t-1} + \phi_1 \varepsilon_t^v + \phi_2 \varepsilon_{t-1}^v \\ &= \phi_1 \varepsilon_t^v + \phi_1(\rho + \phi_2) \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-1-j}^v\end{aligned}$$

where $\phi_1 = 1 - \rho_v(\phi_\pi \psi_\pi + \phi_y \psi_y)$ and $\phi_2 = -\frac{\rho_v(1-\phi_\pi\psi_\pi - \phi_y\psi_y)}{1-\rho_v(\phi_\pi\psi_\pi+\phi_y\psi_y)}$. However, a standard parameterization of the model suggests that such serial correlation is small, in the sense that $\phi_1(\rho + \phi_2) < 0.1$. To be on the safe side, I estimate (D.3) using GMM with Bartlett-Newey-West robust standard errors, using as instruments four lags of the Effective Fed Funds rate, GDP Deflator, CBO Output Gap, Commodity Price Inflation, Real M2 Growth and the spread between the long-term bond rate and the three-month Treasury Bill rate, following Clarida et al. (2000). Overall, I find no evidence of a fall in persistence ρ in the recent decades.

Dominant Root Estimate of Persistence I confirm this result by obtaining the dominant root of the nominal interest rate over time. As I showed in section 2.2, another procedure to measure persistence is to compute the dominant root of an AR(p) process. I estimate an AR(20) process on nominal interest rates, and obtain the dominant root at each sub-sample period. Our results are reported in Table A.iv. I find no evidence for a change in nominal rates persistence. If anything, I find evidence for a moderate increase in the nominal interest

Variable	1954-2020	1969-2020	1954-1985	1969-1985	1985-2020
Fed Funds rate	0.97	0.97	0.95	0.90	0.95

Table A.iv: Dominant root of an AR(20) for nominal interest rate.

rate dominant root over time.

B.3 Change in the Monetary Stance

The fall in inflation persistence coincides with a structural change in the Fed policy stance around 1985:I, documented in Clarida et al. (2000). Economists generally model the change in the Fed stance around 1985:I as a structural break in the reaction function of Central Banks. In particular, the break is modelled as if the elasticity of nominal rates with respect to inflation, ϕ_π, went from a previous value of 1 to a value closer to 2. That is, as if the Federal Reserve had become more hawkish in the recent decades. To test this, I estimate a standard Taylor rule in which nominal rates are elastic to current inflation and output gap, as the benchmark framework suggests. I test for a structural break in 1985:I, and our findings align with those in the literature,[42]

$$i_t = \alpha_i + \phi_\pi \pi_t + \phi_{\pi,*} \pi_t \mathbb{1}_{\{t \geq t^*\}} + \phi_y \tilde{y}_t + \phi_{y,*} \tilde{y}_t \mathbb{1}_{\{t \geq t^*\}} + v_t \tag{B.1}$$

using Bartlett-Newey-West standard errors that take into consideration serially correlated residuals, which the theoretical framework suggests. I report our results in table A.v. Indeed, our results show that prior to 1985 the elasticity with respect to inflation was around 1.32, and increased to 2.28 in the Great Moderation, close to the findings by Clarida et al. (2000). In the following subsection I link this structural change with the dynamics of inflation produced in the NK model. As I will show, the increase in ϕ_π will effectively reduce inflation volatility but will have no effect on persistence.

B.4 Empirical Evidence on Information Frictions

Rolling Sample Regression I obtain a rolling-sample estimate version of (2.2). Figure A.6 plots the rolling estimate $\beta_{CG,t}$ over time. The figure suggests that information frictions

42. To estimate the Taylor rule I rely on GMM methods, using four lags of the Effective Fed Funds rate, GDP Deflator, CBO Output Gap, Commodity Price Inflation, Real M2 Growth and the spread between the long-term bond rate and the three-month Treasury Bill rate as instruments. The standard NK model incorporates inertia in the Taylor rule via the AR(1) component v_t instead of including lags of the nominal interest rate on the right-hand side of (B.1), which allows us to obtain a closed-form solution of the model.

	(1) Taylor Rule	(2) Break
π_t	1.154^{***}	1.323^{***}
	(0.112)	(0.140)
$\tilde{y}_t$	0.353^{***}	0.309^{**}
	(0.121)	(0.128)
$\pi_t \times \mathbb{1}_{\{t \geq t^*\}}$		0.958^{***}
		(0.284)
Constant	1.518^{***}	-0.517
	(0.442)	(0.844)
Observations	204	204

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.v: Regression table

were reduced after the 1980s, with a smaller local peak in the late 2000s, which coincides with the local peak in inflation persistence in Figure A.3.

Time-Varying Parameter Autorregression I use the following representation of the time-varying parameter regression model

$$\pi_{t+3,t} - \overline{\mathbb{E}}_t \pi_{t+3,t} = \beta_{CG,t}(\overline{\mathbb{E}}_t \pi_{t+3,t} - \overline{\mathbb{E}}_{t-1} \pi_{t+3,t}) + u_t, \quad u_t \sim \mathcal{N}(0, \sigma_u^2)$$

where the time-varying persistence coefficient is assumed to follow a random walk

$$\beta_{CG,t+1} = \beta_{CG,t} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma_\epsilon)$$

The model is estimated using Bayesian methods. Our prior selection is standard, following Nakajima (2011), using the invert Wishart and invert Gamma distributions

$$\beta_{CG,1} \sim \mathcal{N}(0, 10 \times I), \quad \sigma_u \sim \mathcal{IG}(2, 0.02), \quad \Sigma_\epsilon \sim \mathcal{IW}(4, 40 \times I)$$

I plot the estimated $\beta_{CG,t}$ with 95% confidence bands in Figure A.7. After the break in the mid-1980s the estimated values are not statistically significant, unable to reject the FIRE assumption.

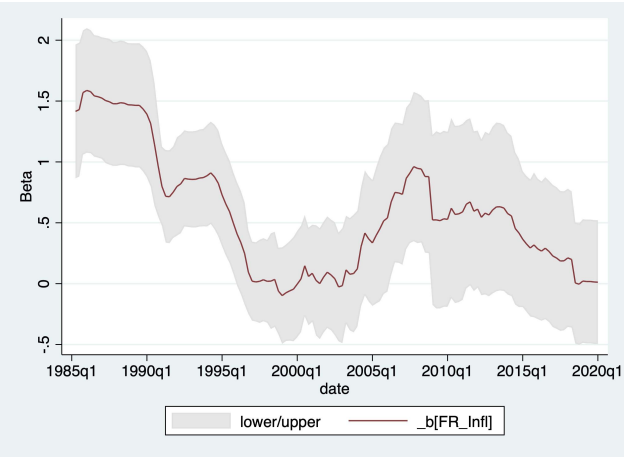


Figure A.6: Time-varying $\beta_{CG,t}$ in the CG regression (2.2) using a 14y window.

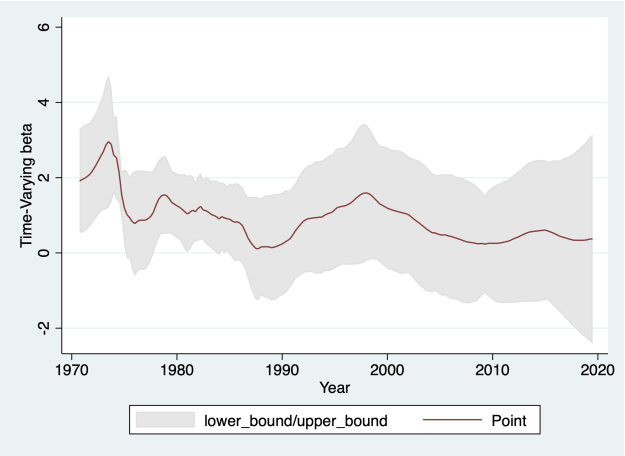


Figure A.7: Time-varying Coibion and Gorodnichenko (2015a) regression.

Forecast Error response to Monetary Policy Shocks Under FIRE, ex-post forecast errors should be unpredictable by ex-ante available information. Therefore, the IRF of forecast errors to monetary policy shocks should be insignificant. Coibion and Gorodnichenko (2012) show that forecast errors react to several exogenous shocks to the economy. In order to study if the sensitivity of ex-post forecast errors has changed after the 1985:Q1 structural break, we produce the local projection of Romer and Romer (2004) monetary policy shocks on the average forecast error,

$$error_{t+h} = \beta_h \varepsilon_t^v + \beta_{h*} \varepsilon_t^v \times \mathbb{1}_{t \geq t^*} + \gamma X_t + u_t$$

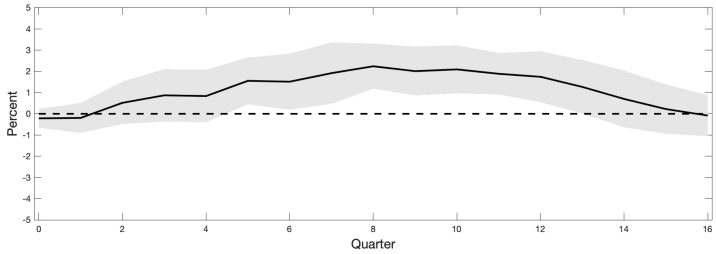
where h denotes the horizon and X_t includes four lags of Romer and Romer (2004) shocks and four lags of forecast errors. We report the implied impulse responses in Figure A.8. We find that the IRF is positive in the pre-1985 period, suggesting that forecasts react less to monetary shocks than the forecasted variable (see Figure A.8a). After 1985, forecast errors do not react to monetary shocks, suggesting that information frictions lessened (see Figure A.8b). I show in Figure A.8c that the difference between the IRFs under the two regimes is significant.

Disagreement I define a measure of "disagreement" as the cross-sectional standard deviation of forecasts at each time,

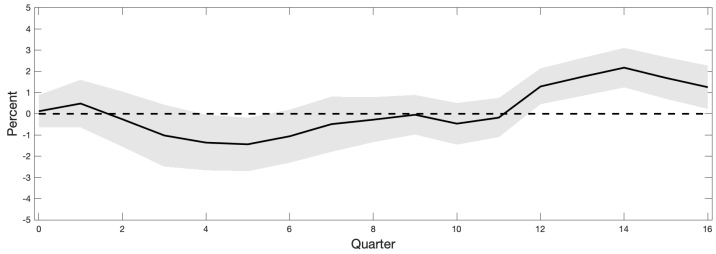
$$disagreement_t = \sigma_i(\mathbb{F}_{it} \pi_{t+3,t})$$

Under the assumption of common complete information, disagreement should be zero since all agents would have observed the same past, their information set would therefore be the same, and their expectation around a future variable should coincide, provided that agents are ex-ante identical. As we observe in Figure A.9, disagreement was large around the 1980s, coinciding with the beginning of the Volcker activism and the lack of public disclosure of the Federal Reserve decisions, and fell dramatically until the 1990s, stabilizing at that level after the 1990s.

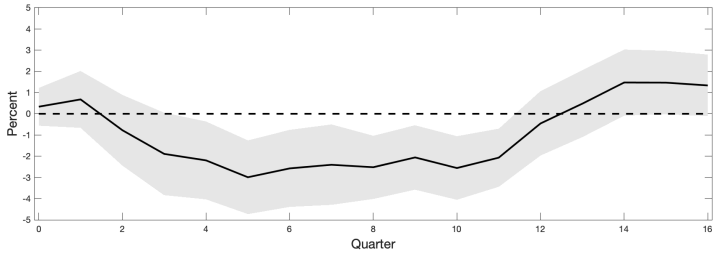
The previous figure dynamics are reminiscent of the inflation dynamics in Figure 1. One could thus argue that, if forecast disagreement depends on the level of inflation, the fall in disagreement would be entirely explained by the fall in inflation. We now show that forecast do not depend on the current level of inflation. First, assuming that inflation follows an



(a) Pre-1985 period.



(b) Post-1985 period.



(c) Change.

Figure A.8: Impulse response function of average forecasts to monetary policy shocks.

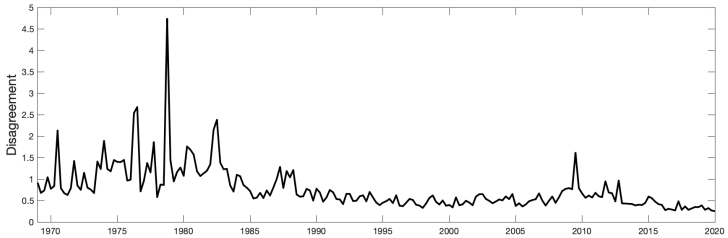


Figure A.9: Cross-sectional volatility of (annual) inflation forecasts at each period.

	Individual forecasts			Average forecast			Error	Error
	AR(1)	AR(2)	AR(3)	AR(1)	AR(2)	AR(3)		
$\mathbb{F}_t\pi_{t+2}$	1.284***	1.435***	1.417***	1.356***	1.870***	1.749***		
	(0.0162)	(0.0476)	(0.0482)	(0.0190)	(0.0707)	(0.0739)		
$\mathbb{F}_t\pi_{t+1}$		−0.232***	−0.0992		−0.775***	−0.390***		
		(0.0652)	(0.0874)		(0.102)	(0.139)		
$\mathbb{F}_t\pi_t$			−0.214***			−0.414***		
			(0.0697)			(0.097)		
$revision_t$							1.220***	
							(0.248)	
$error_{t-1}$								0.881***
								(0.0592)
$\pi_{t-1,t-5}$	0.00705	0.0119	0.0137*	−0.0299**	−0.0182	−0.0169	0.00819	−0.0163
	(0.00909)	(0.00859)	(0.00819)	(0.0124)	(0.0115)	(0.0108)	(0.0340)	(0.0131)
Observations	7,751	7,750	7,750	205	205	205	197	203

HAC robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.vi: Regression table

AR(p) (up to $p = 3$), we regress the individual (average) forecast of the AR(p) process, and we add realized inflation

$$\mathbb{F}_{it}\pi_{t+3} = \rho_1\mathbb{F}_{it}\pi_{t+2} + \rho_2\mathbb{F}_{it}\pi_{t+1} + \rho_3\mathbb{F}_{it}\pi_t + \gamma\pi_{t-1,t-5} + u_t \tag{B.2}$$

$$\overline{\mathbb{F}}_t\pi_{t+3} = \rho_1\overline{\mathbb{F}}_t\pi_{t+2} + \rho_2\overline{\mathbb{F}}_t\pi_{t+1} + \rho_3\overline{\mathbb{F}}_t\pi_t + \gamma\pi_{t-1,t-5} + u_t \tag{B.3}$$

We report our results in Table A.vi. We find in columns 1-3 (columns 4-6) that the lagged inflation coefficient is insignificant in most cases. We then regress the average forecast error on the average forecast revision (column 7) or on the lagged forecast error (column 8), controlling for lagged inflation. We find that realized inflation is insignificant in both cases. We can therefore argue that the fall in cross-sectional forecast volatility fell after the 1985 period as a result of lessened information frictions.

Livingston Survey Using the Livingston survey on firms, I test for a structural break in belief formation around 1985:I. Since the survey is conducted semiannually, I estimate the following structural-break variant of (2.3)

$$\pi_{t+2,t} - \mathbb{E}_t\pi_{t+2,t} = \alpha_{CG} + \left(\beta_{CG} + \beta_{CG*}\mathbb{1}_{\{t\geq t^*\}}\right)\left(\mathbb{E}_t\pi_{t+2,t} - \mathbb{E}_{t-2}\pi_{t+2,t}\right) + u_t \tag{B.4}$$

Our results, reported in the first column in Table A.vii, suggest a strong violation of the FIRE assumption: the measure of information frictions, β_{CG}, is significantly different from zero. Secondly, a significant estimate of β_{CG*} would suggest a break in the information

frictions faced by agents. Our results in the second column in Table III suggest that there is a structural break around the period in which the Fed changed the monetary stance. Our result $\beta_{CG*} < 0$ suggests that agents became *more* more informed about inflation, with individual forecasts relying less on priors and more on news. A t-test under the null that $\beta_{CG} + \beta_{CG,*} = 0$ has an associated p-value of 0.254. I can therefore conclude that information frictions on the CPI vanish, consistent with our findings on CPI persistence in Figure A.3b.

	(1) CG Regression	(2) Structural Break
Revision	0.380^{*}	0.412^{**}
	(0.202)	(0.204)
Revision$\times \mathbb{1}_{\{t \geq t^*\}}$		-0.880^{**}
		(0.414)
Constant	-0.183^{*}	-0.105
	(0.102)	(0.119)
Observations	146	146

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.vii: Regression table

As a second exercise, I estimate the following structural-break variant of (4.6)

$$\pi_{t,t-2} = \left(\alpha_1 + \alpha_{1,*}\mathbb{1}_{\{t \geq t^*\}}\right) \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}_t^f \tilde{y}_{t+k,t+k-2} + \left(\alpha_2 + \alpha_{2,*}\mathbb{1}_{\{t \geq t^*\}}\right) \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}_t^f \pi_{t+k,t+k-2}$$

Since the survey is only conducted semiannually and only asks for 6m and 12m ahead forecasts we only consider the cases $k = 2$ and $k = 4$. Our results suggest no evidence of a structural break in κ once we control for non-standard expectations.

	(1) NKPC	(2) Break Output	(3) Break
$\overline{\mathbb{E}}_t^f \tilde{y}_{t+2,t}$	1.014***	1.402***	1.079**
	(0.262)	(0.438)	(0.418)
$\overline{\mathbb{E}}_t^f \tilde{y}_{t+4,t+2}$	−0.0717	−0.680	−0.354
	(0.335)	(0.553)	(0.533)
$\overline{\mathbb{E}}_t^f \pi_{t+2,t}$	−0.0552	−0.0352	−0.264***
	(0.0652)	(0.0602)	(0.0836)
$\overline{\mathbb{E}}_t^f \pi_{t+4,t+2}$	−0.0375	−0.123	0.237
	(0.151)	(0.147)	(0.180)
$\overline{\mathbb{E}}_t^f \tilde{y}_{t+2,t} \times \mathbb{1}_{\{t \geq t^*\}}$		−0.892*	−0.598
		(0.526)	(0.509)
$\overline{\mathbb{E}}_t^f \tilde{y}_{t+4,t+2} \times \mathbb{1}_{\{t \geq t^*\}}$		0.882	0.555
		(0.662)	(0.641)
$\overline{\mathbb{E}}_t^f \pi_{t+2,t} \times \mathbb{1}_{\{t \geq t^*\}}$			0.303***
			(0.0955)
$\overline{\mathbb{E}}_t^f \pi_{t+4,t+2} \times \mathbb{1}_{\{t \geq t^*\}}$			−0.486**
			(0.191)
Constant	−0.115	0.388	0.479
	(0.250)	(0.398)	(0.460)
Observations	99	99	99

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.viii: Regression table

C Extending Information Frictions to Households

In this section we relax the FIRE assumption on households. We show in Online Appendix F that in such case, the individual household policy function is given by

$$c_{it} = -\frac{\beta}{\sigma}\mathbb{E}_{it}r_t + (1-\beta)\mathbb{E}_{it}\tilde{y}_t + \beta\mathbb{E}_{it}c_{i,t+1}, \quad \text{with } \tilde{y}_t = \int c_{it}\, di \tag{C.1}$$

We still maintain the FIRE assumption on the monetary authority, which is not subject to information frictions. In this case, the model equations are (C.1), (3.5), (3.7) and (3.8).

Information Structure In order to generate heterogeneous beliefs and sticky forecasts, I assume that the information is incomplete and dispersed. Each agent l in group $g \in$ {household, firm} observes a noisy signal x_{lgt} that contains information on the monetary shock v_t, and takes the standard functional form of "outcome plus noise". Formally, signal x_{lgt} is described as

$$x_{lgt} = v_t + \sigma_{gu}u_{lgt}, \quad \text{with } u_{lgt} \sim \mathcal{N}(0,1) \tag{C.2}$$

where signals are agent-specific. This implies that each agent's information set is different, and therefore generates heterogeneous information sets across the population of households and firms. Notice that we allow for heterogeneity in the variance that each of the groups (households and firms) face.

An equilibrium must therefore satisfy the individual-level optimal pricing policy functions (3.5), the individual DIS curve (C.1), the Taylor rule (3.7), and rational expectation formation should be consistent with the exogenous monetary shock process (3.8) and the signal process (C.2).

The following proposition outlines inflation and output gap dynamics.

Proposition 5. *Under noisy information the output gap, price level and inflation dynamics are given by*

$$\boldsymbol{a}_t = A(\vartheta_1, \vartheta_2, \vartheta_3)\boldsymbol{a}_{t-1} + B(\vartheta_1, \vartheta_2, \vartheta_3)v_t \tag{C.3}$$

where $\boldsymbol{a}_t = \begin{bmatrix}\tilde{y}_t & p_t & \pi_t\end{bmatrix}'$ *is a vector containing output, price level and inflation,* $A(\vartheta_1, \vartheta_2, \vartheta_3)$ *is a* 3×3 *matrix and* $B(\vartheta_1, \vartheta_2, \vartheta_3)$ *is a* 3×1 *vector, where* $(\vartheta_1, \vartheta_2, \vartheta_3)$ *are three scalars that are given by the reciprocal of three of the four outside roots of the characteristic polynomial*

of the following matrix[43]

$$\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix}$$

where

$$C_{11}(z) = \left[(z-\beta)(z-\lambda_1)(1-\lambda_1 z) - \left(1-\frac{\lambda_1}{\rho}\right)(1-\rho\lambda_1)\left(1-\beta\left(1+\frac{\phi_y}{\sigma}\right)\right)z^2\right](1-\theta_2 z)$$

$$C_{12}(z) = -(1-\theta)\left(1-\frac{\lambda_1}{\rho}\right)(1-\rho\lambda_1)z\left(\frac{\beta\phi_\pi}{\sigma}z^2 - \frac{\beta(1+\phi_\pi)}{\sigma}z + \frac{\beta}{\sigma}\right)$$

$$C_{21}(z) = -\left(1-\frac{\lambda_2}{\rho}\right)(1-\rho\lambda_2)(1-\theta z)\frac{\kappa\theta}{1-\theta}z^2$$

$$C_{22}(z) = (z-\beta\theta)(z-\lambda_2)(1-\lambda_2 z)(1-\theta z) - (1-\theta)\left(1-\frac{\lambda_2}{\rho}\right)(1-\rho\lambda_2)(1-\beta\theta)z^2$$

with λ_g, $g \in \{1,2\}$ *being the inside root of the following polynomial*

$$\boldsymbol{D}(z) \equiv z^2 - \left(\frac{1}{\rho} + \rho + \frac{\sigma_\epsilon^2}{\rho\sigma_{gu}^2}\right)z + 1$$

Proof. See Appendix A □

In the noisy information framework, inflation is intrinsically persistent and its persistence is governed by the new information-related parameters ϑ_1, ϑ_2 and ϑ_3, as opposed to the benchmark framework in which it is only extrinsically persistent, $A(0,0,0) = \mathbf{0}$. The intuition for this result is simple: inflation is partially determined by expectations (see condition (3.9) under noisy information, or (2.4) under complete information). Under noisy information, expectations are anchored and follow an autoregressive process (see (3.11)), which creates the additional source of anchoring in inflation dynamics, measured by ϑ_1, ϑ_2 and ϑ_3.

Empirical Evidence on Household's Information Frictions There are now two different information parameters to calibrate, since we allow for heterogeneity in information precision by group. In order to calibrate the additional one, we use the Michigan Survey of Consumers' annual forecasts of inflation.[44] Consider the average forecast of annual inflation at time t, $\overline{\mathbb{E}}_t^c \pi_{t+3,t}$, where $\pi_{t+3,t}$ is the inflation between periods $t+3$ and $t-1$. We can

43. The other outside root is always equal to θ and is cancelled out.

44. Each quarter, the University of Michigan surveys 500–1,500 households and asks them about their expectation of price changes over the course of the next year.

think of this object as the action that the average consumer makes. A drawback of this source of expectations data is that it is are only available at a forecasting horizon of one year and therefore revisions in forecasts over identical horizons are not available. Thus, I follow Coibion and Gorodnichenko (2015a) and replace the forecast revision with the change in the year-aheadforecast, yielding the following quasi-revision: $\text{revision}_t \equiv \overline{\mathbb{E}}^c_t \pi_{t+3,t} - \overline{\mathbb{E}}^c_{t-1} \pi_{t+2,t-1}$. The average forecast revision provides information about the average agent annual forecast after the inflow of information between periods t and $t-1$. Recent research (Coibion and Gorodnichenko 2012, 2015a) has documented a positive co-movement between ex-ante average forecast errors and average forecast revisions.[45] Formally, the regression design is

$$\text{forecast error}_t = \alpha_{\text{rev}} + \beta_{\text{rev}} \text{ revision}_t + u_t \tag{C.4}$$

The error term now consists of the rational expectations forecast error and $\beta_{\text{rev}}(\overline{\mathbb{E}}^c_{t-1}\pi_{t-1} - \overline{\mathbb{E}}^c_t \pi_{t+3})$ because forecasts horizons do not overlap. We therefore rely on an IV estimator, using as an instrument the (log) change in the oil price.[46]

Notice that a positive co-movement ($\hat{\beta}_{\text{rev}} > 0$) suggests that positive revisions predict positive forecast errors. That is, after a positive revision of annual inflation forecasts, consumers consistently under-predict inflation. The results, reported in the first column in Table A.ix, suggest a strong violation of the FIRE assumption: the measure of information frictions, β_{rev}, is significantly different from zero. Agents underrevise their forecasts: a positive β_{rev} coefficient suggests that positive revisions predict positive (and larger) forecast errors. In particular, a 1 percentage point revision predicts a 1.012 percentage point forecast error. The average forecast is thus smaller than the realized outcome, which suggests that the forecast revision was too small, or that forecasts react sluggishly.

Following the previous analyses on inflation persistence, I assume that the break date is 1985:Q1. I test for the null of no structural break in inflation dynamics around 1985:Q1.[47] We cannot the null of no break (p-value $= 0.60$). Following a similar structural break analysis as in Section 2.1, I study if there is a change in expectation formation (stickiness) around the same break date. Formally, I test for a structural break in belief formation around 1985:Q1

45. We used the first-release value of annual inflation, since forecasters did not have access to future revisions of the data.

46. Coibion and Gorodnichenko (2015a) argue that oil prices have significant effects on CPI inflation, and therefore are statistically significant predictors of contemporaneous changes in inflation forecasts and can account for an importantshare of their volatility.

47. If we instead are agnostic about the break date(s), the test suggests that there is no such break.

	(1) All Sample	(2) Structural Break
Revision	1.012^{***}	1.706^{*}
	(0.299)	(1.018)
Revision $\times \mathbb{1}_{\{t \geq t^*\}}$		-1.083
		(1.066)
Constant	-0.571^{***}	-0.571^{***}
	(0.181)	(0.180)
Observations	182	182

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.ix: Regression table

by estimating the following structural-break version of (C.4),

$$\text{forecast error}_t = \alpha_{\text{rev}} + \left(\beta_{\text{rev}} + \beta_{\text{rev}*}\mathbb{1}_{\{t \geq t^*\}}\right) \text{revision}_t + u_t \tag{C.5}$$

A significant estimate of $\beta_{\text{rev}*}$ suggests a break in the information frictions. The results in the second column in Table A.ix suggest that there is no structural break around 1985:Q1.

Results I calibrate the two information volatilities σ_{1u} and σ_{2u} to match jointly the empirical evidence on forecast sluggishness in Tables III and A.ix. This results in $\sigma_{1u} = 13.919$ and $\sigma_{2u} = 12.432$ in the pre-1985 sample, and $\sigma_{1u} = 16.566$ and $\sigma_{2u} = 0.015$. In the pre-1985 period, the model-implied inflation first-order autocorrelation is $\rho_{\pi 1} = 0.796$. In the post-1985 period, inflation persistence falls to 0.686. The fall is smaller because the output gap, which is still intrinsically persistent because of households' information frictions, reduces the overall effect of the fall in firm information frictions. Comparing our model results to the empirical analysis in Tables I and II, I find that the noisy information framework can explain around 50% of the point estimate fall.

D Persistence in NK Models

In this section I study the determinants of inflation persistence in a structural macro framework. I show that the empirical findings documented in the previous section present a puzzle in the NK model. I cover a wide range of NK frameworks and show that they cannot explain

the fall in inflation persistence in an empirically consistent manner. Regarding volatility, I show that its fall can be explained via a change in the monetary stance in the post-Volcker era.

In the benchmark NK model, in which agents form rational expectations using complete information, the demand (output gap) and supply side (inflation) dynamics are modeled as two forward-looking stochastic equations, commonly referred to as the Dynamic IS (DIS) and New Keynesian Phillips (NKPC) curves.[48] Nominal interest rates are set by the Central Bank following a reaction function that takes the form of a standard Taylor rule. The Central Bank reacts to excess inflation and output gap and controls an exogenous component, v_t, which follows an independent AR(1) process which innovations are treated as serially uncorrelated monetary policy shocks.

Inserting the Taylor rule (3.7)-(3.8) into the DIS curve (3.6), one can write the model as a system of two first-order stochastic difference equations that can be solved analytically. In particular, inflation dynamics satisfy

$$\pi_t = -\psi_\pi v_t = \rho\pi_{t-1} - \psi_\pi \sigma_\varepsilon \varepsilon_t^v \tag{D.1}$$

where ψ_π satisfies,

$$\psi_\pi = \frac{\kappa}{(1-\rho\beta)[\sigma(1-\rho)+\phi_y] + \kappa(\phi_\pi - \rho)} \tag{D.2}$$

and output gap dynamics are given by $\tilde{y}_t = -\psi_y v_t = \rho\tilde{y}_{t-1} - \psi_y \sigma_\varepsilon \varepsilon_t^v$. Notice that inflation is proportional to the exogenous shock. As a result inflation will inherit its dynamic properties from the exogenous driving force.[49] A final implication is that inflation is only *extrinsically* persistent: its persistence is determined by the v_t AR(1) process' persistence.

In order to explain the fall in inflation persistence and volatility I discuss each causal explanation separately. First, I explore whether there has been a change in the structural shocks affecting the economy. I show that these exogenous forces' dynamics have been remarkably stable since the beginning of the sample. Second, I investigate if a change in the monetary stance around 1985:Q1, for which Clarida et al. (2000) and Lubik and Schorfheide (2004) provide empirical evidence, could have affected inflation dynamics. I show that the

48. The model derivation is relegated to Online Appendix F.

49. One can also notice that the benchmark model predicts that output gap and inflation are equally persistent, and their dynamics will only differ due to the differential monetary policy shock impact effect, captured by ψ_y and ψ_π. Another implication is that the Pearson correlation coefficient between output gap and inflation is equal to 1, an aspect rejected in the data.

change in the monetary stance can indeed explain the fall in volatility but has null or modest effects on persistence. Finally, I explore if changes in intrinsic persistence, generated via backward-looking assumptions on the firm side, have a sizeable effect on persistence. As in the previous case, I show that these have only marginal effects.

D.1 Structural Shocks

I documented in Section 2.1 that inflation persistence and volatility fell in the recent decades. The NK model suggests that such fall is inherited from a fall in the persistence of the monetary policy shock process. I now seek to find evidence on the time-varying properties of such persistence.

Persistence The challenge that the econometrician faces is that she does not have an empirical proxy for v_t. The monetary policy shocks estimated by the literature are not serially correlated, and are therefore a better picture of the monetary policy shock ε_t^v.[50,51] However, one can use the model properties and rewrite the Taylor rule (3.7) using the AR(1) properties of (3.8), as

$$i_t = \rho i_{t-1} + (\phi_\pi \pi_t + \phi_y y_t) - \rho (\phi_\pi \pi_{t-1} + \phi_y y_{t-1}) + \sigma_\varepsilon \varepsilon_t^v \tag{D.3}$$

where the error term is the monetary policy shock.[52] Hence, an estimate of the first-order autoregressive coefficient in (D.3) identifies the monetary policy shock process persistence.[53] I present here the structural break analysis and leave for Appendix B.2.1 the robustness analysis. I test for a potential tructural break in the persistence of the nominal interest rate process, described by (D.3), around 1985:Q1. I do this in two different ways. First, I use an unrestricted GMM and estimate

$$i_t = \alpha_i + \alpha_{i,*} \mathbb{1}_{\{t \geq t^*\}} + \rho_i i_{t-1} + \rho_{i,*} i_{t-1} \mathbb{1}_{\{t \geq t^*\}} + \boldsymbol{\gamma} \boldsymbol{X}_t + u_t$$

50. In fact, the process v_t is a model device engineered to produce inertia yet still allowing us to obtain a closed-form solution. If inertia is directly introduced in the nominal interest rate equation, I would not be able to obtain the closed-form solution (D.1) since the system would also feature a backward-looking term whose coefficients would depend on the roots of a quadratic polynomial.

51. For example, Romer and Romer (2004) use the cumulative sum of their estimated monetary policy shocks to derive the IRFs.

52. Using the lag operator, I can write the monetary policy shock process (3.8) as $v_t = (1 - \rho L)^{-1} \varepsilon_t^v$. Introducing this last expression into (3.7), multiplying by $(1 - \rho L)$ and rearranging terms, I obtain (D.3).

53. Our measure of the nominal rate will be the effective Fed Funds rate (EFFR), calculated as a volume-weighted median of overnight federal funds transactions, and is available at daily frequency. I use the quarterly frequency series.

Panel A	(1) Unrestricted GMM	(2) Restricted GMM
i_{t-1}	0.939^{***}	0.931^{***}
	(0.0448)	(0.0365)
$i_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$	-0.00261	-0.0537
	(0.0591)	(0.0632)
Constant	0.305	0.851^{**}
	(0.473)	(0.373)
Constant$\times \mathbb{1}_{\{t \geq t^*\}}$	-0.123	-0.813
	(0.436)	(0.559)
Observations	203	203
Panel B	Romer & Romer	
	Pre 1985	Post 1985
Standard deviation	0.286	0.0923

HAC robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.x: Regression table

where $\boldsymbol{X}_t$ is a set of control variables that includes current and lagged output gap and inflation.[54] I report our results in the first two columns of Table A.x Panel A. There is no evidence for a decrease in nominal interest rate persistence (and thus, monetary shock process persistence) over time. Notice however that monetary shock persistence plays a dual role in (D.3), since it also affects lagged output and inflation. As a robustness check, I estimate the structural break version of (D.3) using a restricted-coefficient GMM, reported in the last two columns in Table A.x Panel A. Our findings are similar.

This set of results is inconsistent with the NK model, since the model suggests that the empirically documented fall in inflation persistence can only be explained by an *identical* fall in nominal interest rates persistence.

Additional Structural shocks In the model studied above I only considered monetary policy shocks, but it could be the case that other relevant shocks have lost persistence in recent decades and could thus explain the fall in inflation persistence. I additionally consider

54. The instrument set includes four lags of the Effective Fed Funds rate, GDP Deflator, CBO Output Gap, labor share, Commodity Price Inflation, Real M2 Growth and the spread between the long-term bond rate and the three-month Treasury Bill rate.

Panel A: Model	Model	
Persistence	Pre 1985	Post 1985
Monetary	0.50	0.50
Add technology & cost-push	0.80	0.80
Panel B: Data		
First-Order Autocorrelation	Pre 1985	Post 1985
Technology Shocks	0.934	0.980
Cost-push Shocks	0.933	0.913

Table A.xi: Summary

demand (technology) and supply (cost-push) shocks. In this case inflation dynamics follow

$$\pi_t = \psi_{\pi v} v_t + \psi_{\pi a} a_t + \psi_{\pi u} u_t \tag{D.4}$$

where a_t is the technology shock, u_t is the cost-push shock, $\psi_{\pi x}$ for $x \in \{v, a, u\}$ are scalars that depend on model parameters, defined in Online Appendix G, and shock processes follow respective AR(1) processes $x_t = \rho_x x_{t-1} + \varepsilon_t^x$. Using different measures of technology shocks from Fernald (2014), Francis et al. (2014), and Justiniano et al. (2011) and cost-push shocks from Nekarda and Ramey (2010), I show in Online Appendix G that there is no empirical evidence for a fall in their persistence. Additionally, I find that an increase in ϕ_π from 1 to 2, as the one documented by Clarida et al. (2000), can only generate a fall of 0.003% in the first-order autocorrelation. Therefore, I can rule out this explanation.

D.2 Monetary Stance

We now consider exogenous changes in the reaction function of the monetary authority. Let us first consider the benchmark framework, with inflation dynamics described by (D.1). I already argued that changes in the policy rule do not affect inflation persistence. Let us now consider extensions of the benchmark model that could explain the fall in inflation persistence. I begin by considering a hypothetical change in monetary policy, conducted via the Taylor rule (3.7)-(3.8). The previous literature has considered the possibility of the Fed conducting a passive monetary policy before 1985, which in the lens of the theory would lead to multiplicity of equilibria. For example, Clarida et al. (2000) document that the inflation coefficient in the Taylor rule was well below one, not satisfying the Taylor principle. Lubik and Schorfheide (2004) estimate a NK model under determinacy and indeterminacy, and argue that monetary policy after 1982 is consistent with determinacy, whereas the pre-

Volcker policy is not. I study if this change in the monetary stance could have affected inflation persistence. I find that inflation dynamics are more persistent in the indeterminacy region, with an autocorrelation of 0.643, falling to 0.5 in the determinacy region after the mid 1980s. This could explain more than 50% of the overall fall in inflation persistence. Another interesting result is that, even in the case of multiple equilibria arising from non-fundamental sunspot shocks, the first-order autocorrelation coefficient is unique.

The second extension that I inspect is optimal monetary policy under discretion. I show that an increase in ϕ_π can be micro-founded through a change in the monetary stance in which the central bank follows a Taylor rule in the pre-1985 period, while it follows optimal monetary policy under discretion in the post-1985 period. In such case, inflation dynamics follow (D.4) in the pre-1985 period, and $\pi_t = \rho_u \pi_{t-1} + \psi_d \varepsilon_t^u$ in the post-1985 period, where ψ_d is a positive scalar that depends on deep parameters and inflation persistence is inherited from the cost-push shock. Compared to the pre-1985 dynamics, described by (D.4), there is no significant change in inflation persistence: in the pre-period, model persistence is around 0.80,[55] while in the post-period persistence is around 0.80.[56] Therefore, such change in the policy stance would have generated an *increase* in inflation persistence, which rules out this explanation.

Consider the benchmark NK model with optimal monetary policy under commitment. Under commitment, the monetary authority can credibly control households' and firms' expectations. In this framework, inflation dynamics are given by $\pi_t = \rho_c \pi_{t-1} + \psi_c \Delta u_t$, where ρ_c and ψ_c are positive scalars that depend on deep parameters, $\Delta u_t \equiv u_t - u_{t-1}$ is the exogenous cost-push shock process, with ρ_c governing inflation intrinsic persistence. Using a standard parameterization I find that $\rho_c = 0.310$, which suggests that this framework, although it produces an excessive fall in inflation persistence, could explain its fall. Its main drawback is that its *implied* Taylor rule in the post-1985 period would require an increase in ϕ_π from 1 to 4.5, as I show in Online Appendix G, which is inconsistent with the documented evidence in table A.x Panel A.

D.3 Intrinsic Persistence

The main reason for the failure in explaining the change in the dynamics in the benchmark NK model is that the endogenous outcome variables, output gap and inflation, are propor-

55. Measured by the first-order autocorrelation of (D.4).

56. The estimated persistence of cost-push shocks, ρ_u, is constant throughout both periods, as I document in Table OA.4.

	Model	
Persistence	Pre 1985	Post 1985
Indeterminacy	0.643	0.5
Discretion	0.799	0.800
Commitment	0.799	0.400

Table A.xii: Summary

tional to the monetary policy shock process and thus inherit its dynamics. This is a result of having a pure forward-looking model, which direct consequence is that endogenous variables are not intrinsically persistent, and its persistence is simply inherited from the exogenous driving force and unaffected by changes in the monetary stance. I therefore enlarge the standard NK model to accommodate a backward-looking dimension in the following discussed extensions, including a lagged term in the system of equations.

I consider a backward-looking inflation framework, "micro-founded" through price indexation. In this framework, a restricted firm resets its price (partially) indexed to past inflation, which generates anchoring in aggregate inflation dynamics. In such framework, inflation dynamics are given by $\pi_t = \rho_\omega \pi_{t-1} + \psi_\omega \upsilon_t$. In this framework inflation intrinsic persistence is increasing in the degree of price indexation ω, as I show in Online Appendix G. A fall in the degree of indexation could explain the fall in inflation persistence. However, the parameterization of such parameter is not a clear one. Price indexation implies that every price is changed every period, and therefore one could not identify the Calvo restricted firms in the data and estimate ω. As a result, the parameter is usually estimated using aggregate data and trying to match the anchoring of the inflation dynamics, and its estimate will therefore depend on the additional model equations. Christiano et al. (2005) assume $\omega = 1$. Smets and Wouters (2007) estimate a value of $\omega = 0.21$ trying to match aggregate anchoring in inflation dynamics. It is hard to justify a particular micro estimate for ω, since it is unobservable in the micro data.[57] A counterfactual prediction in this framework is that all prices are changed in every period, in contradiction with the empirical findings in Bils and Klenow (2004) and Nakamura and Steinsson (2008). As a result, one cannot credibly claim that ω is the causant of the fall in inflation persistence, since it needs to be identified from the macro aggregate data, which makes unfeasible to identify ω and the true inflation persistence separately. Finally, I find that a change in the monetary policy stance has now a significant effect on inflation persistence: a change of ϕ_π from 1 to 2 produces a fall in the

57. One would need to identify the firms that were not hit by the Calvo fairy in a given period, yet they change their price.

	Model	
Persistence	Pre 1985	Post 1985
Price indexation	0.90	0.87
Trend inflation	0.91	0.84

Table A.xiii: Summary

first-order autocorrelation of inflation from around 0.895 to 0.865. However, is not enough to produce the effect that I observe in the data.

Our last extension is to include trend inflation, for which the literature has documented a fall from 4% in the 1947-1985 period to 2% afterwards (see e.g., Ascari and Sbordone 2014; Stock and Watson 2007). Differently from the standard environment, I log-linearize the model equations around a steady state with positive trend inflation, which I assume constant within eras. Augmenting the model with trend inflation creates intrinsic persistence in the inflation dynamics through relative price dispersion, which is a backward-looking variable that has no first-order effects in the benchmark NK model. Inflation dynamics are now given by $\pi_t = \rho_{\bar{\pi},1}\pi_{t-1} + \rho_{\bar{\pi},2}\pi_{t-2} + \psi_{\bar{\pi},1}v_t + \psi_{\bar{\pi},2}v_{t-1}$, where persistence is increasing in the level of trend inflation. I therefore investigate if the documented fall in trend inflation, coupled with the already discussed change in the monetary stance, can explain the fall in inflation persistence. Although in the correct direction, I find that the fall in trend inflation and the increase in the Taylor rule coefficients produce a small decrease in intrinsic persistence, from 0.91 to 0.84.

E History of Fed's Gradual Transparency

Fed's actions have become more transparent over time. Before 1967 the FOMC only announced policy decisions once a year in the Annual Report. The report also included the Memoranda of Discussion (MOD) containing the minutes of the meeting, released with a 5 year lag since 1935. In 1967, the FOMC decided to release the directive in the PR, 90 days after the decision. The rationale for maintaining a delay was that earlier disclosure would interfere with central bank best practice due to political pressure, both from the Administration and from the Congress. In a letter from Chairman Burns to Senator Proxmire on August 1972, Burns enumerated six reasons for deferment of availability. Among them, Burns argued that earlier disclosure could interfere with the execution of policies, permit speculators to gain unfair profits by trading in securities, foreign exchange, etc., result in unwarranted disturbances in the asset market, or affect transactions with foreign govern-

ments or banks. In the same letter Burns hypothesised with reducing the delay shorter than 90 days, although stressing that a few hours/days delay would harm the Fed.

In March 1975 David R. Merril, a student at Georgetown University, requested *current* MOD to be disclosed based on the Freedom of Information Act (FOIA). Congressman Patman supported this initiative, and officially asked Chairman Burns for the unedited MOD from the period 1971-1974. Burns declined to comply with the request.[58] At the same time, the FOMC formed a subcommittee on the matter, which suggested to cut back substantially on details about the members' forecasts and to allow each member to edit the minutes, but discouraged eliminating the MODs. In May 1976, concerned about the chance of premature disclosure, the FOMC discontinued the MOD arguing that it had not been a useful tool.[59,60] The decision increased the ire of several critics of the Fed. In the coming years the Congress took several actions to protect the premature release of the minutes, in order to convince the Fed to reinstate the MOD, with no success. Contemporaneously to these events, in May 1976 the PR increased its length (expanded to include short-run and long-run members' forecasts) and reduced the delay to 45 days, shortly after the next (monthly) meeting.

Merrill's lawsuit included the request for an immediate release of the directive (the Fed decision). On November 1977 the Court of Appeals for the District of Columbia ruled in Merrill's favor on this regard. In January 1978, Burns asked Senator Proxmire for legislative relief from the requirement. Finally, in June 1979 the Supreme Court ruled in the FOMC favor.

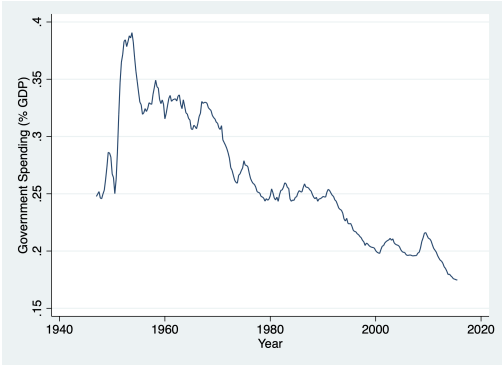
Between 1976 and 1993 the information contained in the PR was significantly enlarged, without further changes in the announcement delay. In November 1977 the *Federal Reserve Reform Act* officially entitled the Fed with 3 objectives: maximum employment, stable prices and moderate long-term interest rates. In July 1979, the first individual macroeconomic forecasts on (annual) real GNP growth, GNP inflation and unemployment from FOMC members were made available. During this period, the Fed was widely criticised for the rise in inflation (see Figure 1). The FOMC stressed in their communication that the increase in inflation was due to excessive fiscal policy stimulus (see Figure A.10a) and the cost-push shock on real wages coming from the increased worker unionization (see Figure A.10b).

From October 1979 to November 1989 the policy instrument changed from the fed funds

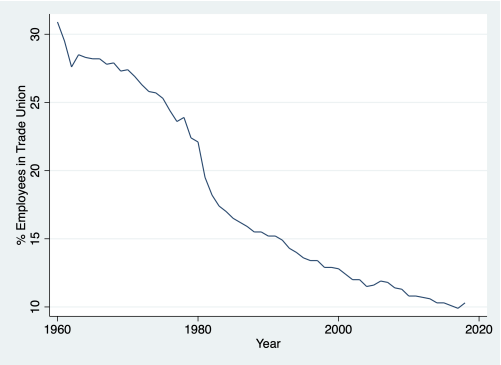
58. The letter exchange is available at Lindsey (2003), pp. 11-15.

59. Robert P. Black, former president of the Richmond Fed that served at the FOMC, explained years later that "I did it for the fear that Congress would request access quite promptly" (see Lindsey (2003), p. 22).

60. Whether meetings were still recorded was unclear to the public, until Chairman Greenspan revealed their existence in October 1993, causing a stir.



(a) Real government spending as a share of real GDP.



(b) Percentage of workers that are members of a Trade Union.

Figure A.10: Time series.

rate to non-borrowed reserves (M1, until Fall 1982) and borrowed reserves (M2 and M3, thereafter), respectively. In the early 1980s the Fed had not stablished an inflation target yet. Instead, the focus was on stabilizing monetary aggregates, M1 growth in particular. However, frequent and volatile changes in money demand made it particularly challenging for the Fed to deliver stable monetary aggregates. The aspects of these operational procedures were not explained to the public during 1982.

The "tilt" (predisposition or likelihood regarding possible future action) was introduced in the PR in November 1983. Between March 1985 and December 1991 the Fed introduced the "ranking of policy factors", which after each meeting ranked aggregate macro variables in importance, signaling priorities with regard to possible future adjustments. During this period the FOMC members started discussing internally the possibility of reducing the delay of announcements. An internal report from November 1982 summarizes the benefits, calling for democratic public institutions, reducing the criticism due to excessive secrecy, and the induced misallocation of resources by firms, somehow forced to hire "Fed watchers". Yet, the cons, which remained similar as those expressed in 1972. In fact, Chairman Volcker defended the Fed's translucent policy in two letters to Representative Fauntroy in August 1984 and Senator Mattingly in July 1985.

Until then, the FOMC had been successful in convincing politicians and the judicial system that its secrecy was grounded in a purely economic rationale, and was not the result of an arbitrary decision. The first critique from the academic profession came from Goodfriend (1986), which argued that opaqueness reduces the power of monetary policy by distorting agents' reactions. Cukierman and Meltzer (1986) formalize a theoretical framework in which

credibility and reputation induce rich dynamics around a low-inflation steady state. Blinder (2000) and Bernanke et al. (1999) stressed the benefits of a more transparent policy, such as inflation targeting. Faust and Svensson (2001) build a framework in which the Central Bank cares about its reputation, and identify a potential conflict between society and the Central Bank: the general public wants full transparency, while the Central Bank prefers minimal transparency. Faust and Svensson (2002) extend their results by endogeneizing the choice of transparency and the degree of control that the Central Bank has.

After the successful disinflation episode in the mid 1980s the Fed gained reputation, not fearing criticism of further tightening in the policy stance. As a result the FOMC was subject to little political interference, which together with the criticism coming from the academic profession led them to increase transparency. The minutes, a revised transcript of the discussions during the meeting, were reintroduced into the PR in March 1993 under Chairman Greenspan. In 1994 the FOMC introduced the immediate release of the PR after a meeting if there had been a decision, coupled with an immediate release of the "tilt" since 1999. Since January 2000 there is an immediate announcement and press conference after each meeting, regardless of the decision.

Online Appendix

F Model Derivation

F.1 Derivation of the General New Keynesian Model

F.1.1 Households

There is a continuum of infinitely-lived, ex-ante identical households indexed by $i \in \mathcal{I}_h = [0,1]$ seeking to maximize

$$\mathbb{E}_{i0} \sum_{t=0}^{\infty} \beta^t U(C_{it}, N_{it}) \tag{F.1}$$

where utility takes a standard CRRA shape $U(C,N) = \frac{C^{1-\sigma}}{1-\sigma} - \frac{N^{1+\varphi}}{1+\varphi}$. Notice that I relax the benchmark framework and assume that households might differ in their beliefs and their expectation formation. Furthermore, the consumption index C_{it} is given by

$$C_{it} = \left(\int_{\mathcal{I}_f} C_{ijt}^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}$$

with C_{ijt} denoting the quantity of good j consumed by household i in period t, and ϵ denotes the elasticity between goods. Here I have assumed that each consumption good is indexed by $j \in \mathcal{I}_f = [0,1]$. Given the different good varieties, the household must decide how to optimally allocate its limited expenditure on each good j. A cost-minimization problem yields

$$C_{ijt} = \left(\frac{P_{jt}}{P_t} \right)^{-\epsilon} C_{it} \tag{F.2}$$

where the aggregate price index is defined as $P_t \equiv \left(\int_{\mathcal{I}_f} P_{jt}^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}$. Using the above conditions, one can show that

$$\int_{\mathcal{I}_f} P_{jt} C_{ijt} \, dj = P_t C_{it}$$

I can now state the household-level budget constraint. In real terms, households decide how much to consume, work and save subject to the following restriction

$$C_{it} + B_{it} = R_{t-1} B_{i,t-1} + W_t^r N_{it} + D_t \tag{F.3}$$

where N_{it} denotes employment (or hours worked) by household i, B_{it} denotes savings (or bond purchases) by household i, R_{t-1} denotes the gross real return on savings, W_t^r denotes the real wage at time t, and D_t denotes dividends received from the profits produced by firms. The optimality conditions from the household problem satisfy

$$\begin{aligned} C_{it}^{-\sigma} &= \beta \mathbb{E}_{it} \left(R_t C_{i,t+1}^{-\sigma} \right) \\ C_{it}^{\sigma} N_{it}^{\varphi} &= \mathbb{E}_{it} W_t^r \end{aligned}$$

Let us now focus on the budget constraint. Define $A_{it} = R_{t-1} B_{i,t-1}$ as consumer i's initial asset position in period t. Rewrite (F.3) at $t+1$

$$C_{it+1} + B_{it+1} = R_t B_{i,t} + W_{t+1}^r N_{it+1} + D_{t+1} \tag{F.4}$$

Combining (F.3) and (F.4) I can write

$$C_{it} + (C_{it+1} + B_{it+1}) R_t^{-1} = A_{it} + W_t^r N_{it} + D_t + (W_{t+1}^r N_{it+1} + D_{t+1}) R_t^{-1}$$

Doing this until $t \to \infty$ I obtain

$$\sum_{k=0}^{\infty} \prod_{j=1}^{k} \frac{1}{R_{t+j-1}} C_{it+k} = A_{it} + \sum_{k=0}^{\infty} \prod_{j=1}^{k} \frac{1}{R_{t+j-1}} (W_{t+k}^r N_{it+k} + D_{t+k})$$

Log-linearizing the above condition around a zero inflation steady-state I obtain

$$\sum_{k=0}^{\infty} \beta^k c_{it+k} = a_{it} + \Omega_i \sum_{k=0}^{\infty} \beta^k (w_{t+k}^r + n_{it+k}) + (1 - \Omega_i) \sum_{k=0}^{\infty} \beta^k d_{t+k} \tag{F.5}$$

where a lower case letter denotes the log deviation from steady state, i.e., $x_t = \log X_t - \log X$, except for the initial asset position, defined as $a_{it} = A_{it}/C_i$; and Ω_i denotes the labor income share for household i.

The optimal intratemporal labor supply condition can be log-linearized to

$$\mathbb{E}_{it} w_t^r = \sigma c_{it} + \varphi n_{it} \tag{F.6}$$

and the intertemporal Euler condition can be log-linearized to

$$c_{it} = -\frac{1}{\sigma}\mathbb{E}_{it}r_t + \mathbb{E}_{it}c_{it+1} \tag{F.7}$$

where I define the ex-post real interest rate as $r_t = i_t - \pi_{t+1}$.

I want to obtain the optimal expenditure of household i in period t as a function of the current a future expected wages, dividends and real interest rates. Using (F.6) and taking expectations, I can rearrange (F.5) as

$$\begin{aligned}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}c_{it+k} &= a_{it} + \Omega_i\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left(\frac{1+\varphi}{\varphi}w^r_{t+k} - \frac{\sigma}{\varphi}c_{it+k}\right) + (1-\Omega_i)\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}d_{t+k} \\ &= \frac{\varphi}{\varphi+\sigma\Omega_i}a_{it} + \sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left[\frac{\Omega_i(1+\varphi)}{\varphi+\sigma\Omega_i}w^r_{t+k} + \frac{(1-\Omega_i)\varphi}{\varphi+\sigma\Omega_i}d_{t+k}\right]\end{aligned} \tag{F.8}$$

Let us now focus on the left-hand side. Taking individual expectations, I can rewrite it as $\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}c_{it+k}$. Keeping this aside, I can rearrange (F.7) as

$$\mathbb{E}_{it}c_{it+1} = c_{it} + \frac{1}{\sigma}\mathbb{E}_{it}r_t$$

Iterating (F.7) one period forward, I can similarly write

$$\mathbb{E}_{it}c_{it+2} = c_{it} + \frac{1}{\sigma}\mathbb{E}_{it}(r_t + r_{t+1})$$

and, for a general k,

$$\mathbb{E}_{it}c_{it+k} = c_{it} + \frac{1}{\sigma}\sum_{j=0}^{k}\mathbb{E}_{it}r_{t+j}$$

That is, I can write

$$\begin{aligned}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}c_{it+k} &= \sum_{k=0}^{\infty}\beta^k c_{it} + \frac{1}{\sigma}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\beta^k\mathbb{E}_{it}r_{t+j} \\ &= \frac{1}{1-\beta}c_{it} + \frac{\beta}{\sigma(1-\beta)}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}r_{t+k}\end{aligned}$$

Inserting this last condition into (F.8), I can write

$$c_{it} = -\frac{\beta}{\sigma}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}r_{t+k} + \frac{\varphi(1-\beta)}{\varphi+\sigma\Omega}a_{it} + \sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left[\frac{\Omega_i(1+\varphi)(1-\beta)}{\varphi+\sigma\Omega}w^r_{t+k} + \frac{(1-\Omega_i)\varphi(1-\beta)}{\varphi+\sigma\Omega}d_{t+k}\right]$$

Aggregating, using the fact that assets are in zero net supply, $\int_{\mathcal{I}_h} a_{it}\,di = a_t = 0$,

$$c_t = -\frac{\beta}{\sigma}\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}^h_t r_{t+k} + \sum_{k=0}^{\infty}\beta^k\left[\frac{\Omega(1+\varphi)(1-\beta)}{\varphi+\sigma\Omega}\overline{\mathbb{E}}^h_t w^r_{t+k} + \frac{(1-\Omega)\varphi(1-\beta)}{\varphi+\sigma\Omega}\overline{\mathbb{E}}^h_t d_{t+k}\right] \quad \text{(F.9)}$$

where $\overline{\mathbb{E}}^h_t(\cdot) = \int_{\mathcal{I}_c}\mathbb{E}_{it}(\cdot)\,di$ is the average household expectation operator in period t.

F.1.2 Firms

As in the household sector, I assume a continuum of firms indexed by $j \in \mathcal{I}_f = [0,1]$. Each firm is a monopolist producing a differentiated intermediate-good variety, producing output Y_{jt} and setting nominal price P_{jt} and making real profit D_{jt}. Technology is represented by the production function

$$Y_{jt} = A_t N_{jt}^{1-\alpha} \quad \text{(F.10)}$$

where A_t is the level of technology, common to all firms, which evolves according to

$$a_t = \rho_a a_{t-1} + \varepsilon^a_t \quad \text{(F.11)}$$

where $\varepsilon^a_t \sim \mathcal{N}(0,\sigma^2_a)$.

Aggregate Price Dynamics As in the benchmark NK model, price rigidities take the form of Calvo-lottery friction. At every period, each firm is able to reset their price with probability $(1-\theta)$, independent of the time of the last price change. That is, only a measure $(1-\theta)$ of firms is able to reset their prices in a given period, and the average duration of a price is given by $1/(1-\theta)$. Such environment implies that aggregate price dynamics are given (in log-linear terms) by

$$\pi_t = \int_{\mathcal{I}_f}\pi_{jt}\,dj = (1-\theta)\left[\int_{\mathcal{I}_f}p^*_{jt}\,dj - p_{t-1}\right] = (1-\theta)\left(p^*_t - p_{t-1}\right) \quad \text{(F.12)}$$

Optimal Price Setting A firm re-optimizing in period t will choose the price P^*_{jt} that maximizes the current market value of the profits generated while the price remains effective. Formally,

$$P^*_{jt} = \arg\max_{P_{jt}} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_{jt} \left\{ \Lambda_{t,t+k} \frac{1}{P_{t+k}} \left[P_{jt} Y_{j,t+k|t} - \mathcal{C}_{t+k}(Y_{j,t+j|t}) \right] \right\}$$

subject to the sequence of demand schedules

$$Y_{j,t+k|t} = \left(\frac{P_{jt}}{P_{t+k}} \right)^{-\epsilon} C_{t+k}$$

where $\Lambda_{t,t+k} \equiv \beta^k \left(\frac{C_{t+k}}{C_t} \right)^{-\sigma}$ is the stochastic discount factor, $\mathcal{C}_t(\cdot)$ is the (nominal) cost function, and $Y_{j,t+k|t}$ denotes output in period $t+k$ for a firm j that last reset its price in period t. The First-Order Condition is

$$\sum_{k=0}^{\infty} \theta^k \mathbb{E}_{jt} \left[\Lambda_{t,t+k} Y_{j,t+k|t} \frac{1}{P_{t+k}} \left(P^*_{jt} - \mathcal{M} \Psi_{j,t+k|t} \right) \right] = 0$$

where $\Psi_{j,t+k|t} \equiv \mathcal{C}'_{t+k}(Y_{j,t+j|t})$ denotes the (nominal) marginal cost for firm j, and $\mathcal{M} = \frac{\epsilon}{\epsilon - 1}$. Log-linearizing around the zero inflation steady-state, I obtain the familiar price-setting rule

$$p^*_{jt} = (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt} \left(\psi_{j,t+k|t} + \mu \right) \tag{F.13}$$

where $\psi_{j,t+k|t} = \log \Psi_{j,t+k|t}$ and $\mu = \log \mathcal{M}$.

F.1.3 Equilibrium

Market clearing in the goods market implies that $Y_{jt} = C_{jt} = \int_{\mathcal{I}_h} C_{ijt}\, di$ for each j good/firm. Aggregating across firms, I obtain the aggregate market clearing condition: since assets are in zero net supply and there is no capital, investment, government consumption nor net exports, production equals consumption:

$$\int_{\mathcal{I}_f} Y_{jt}\, dj = \int_{\mathcal{I}_h} \int_{\mathcal{I}_f} C_{ijt}\, dj\, di \implies Y_t = C_t$$

Aggregate employment is given by the sum of employment across firms, and must meet

aggregate labor supply

$$N_t = \int_{\mathcal{I}_h} N_{it}\, di = \int_{\mathcal{I}_f} N_{jt}\, dj$$

Using the production function (F.10) and (F.2) together with goods market clearing

$$N_t = \int_{\mathcal{I}_f} \left(\frac{Y_{jt}}{A_t}\right)^{\frac{1}{1-\alpha}} dj = \left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}} \int_{\mathcal{I}_f} \left(\frac{P_{jt}}{P_t}\right)^{-\frac{\epsilon}{1-\alpha}} dj$$

Log-linearizing the above expression yields to

$$n_t = \frac{1}{1-\alpha}(y_t - a_t) \tag{F.14}$$

The (log) marginal cost for firm j at time $t+k|t$ is

$$\begin{aligned}\psi_{j,t+k|t} &= w_{t+k} - mpn_{j,t+k|t} \\ &= w_{t+k} - [a_{t+k} - \alpha n_{j,t+k|t} + \log(1-\alpha)]\end{aligned}$$

where $mpn_{j,t+k|t}$ and $n_{j,t+k|t}$ denote (log) marginal product of labor and (log) employment in period $t+k$ for a firm that last reset its price at time t, respectively.

Let $\psi_t \equiv \int_{\mathcal{I}_f} \psi_{jt}$ denote the (log) average marginal cost. I can then write

$$\psi_t = w_t - [a_t - \alpha n_t + \log(1-\alpha)]$$

Thus, the following relation holds

$$\begin{aligned}\psi_{j,t+k|t} &= \psi_{t+k} + \alpha(n_{jt+k|t} - n_{t+k}) \\ &= \psi_{t+k} + \frac{\alpha}{1-\alpha}(y_{jt+k|t} - y_{t+k}) \\ &= \psi_{t+k} - \frac{\alpha\epsilon}{1-\alpha}(p^*_{jt} - p_{t+k})\end{aligned} \tag{F.15}$$

Introducing (F.15) into (F.13), I can rewrite the firm price-setting condition as

$$p^*_{jt} = (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\left(p_{t+k} - \Theta\hat{\mu}_{t+k}\right)$$

where $\hat{\mu} = \mu_t - \mu$ is the deviation between the average and desired markups, where $\mu_t = -(\psi_t - p_t)$, and $\Theta = \frac{1-\alpha}{1-\alpha+\alpha\epsilon}$.

Individual and Aggregate Phillips curve Suppose that firms observe the aggregate prices up to period $t-1$, p^{t-1}, then I can restate the above condition as

$$p^*_{jt} - p_{t-1} = -(1-\beta\theta)\Theta \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt}\hat{\mu}_{t+k} + \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt}\pi_{t+k}$$

Define the firm-specific inflation rate as $\pi_{jt} = (1-\theta)(p^*_{jt} - p_{t-1})$. Then I can write the above expression as

$$\begin{aligned}
\pi_{jt} &= -(1-\theta)(1-\beta\theta)\Theta \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt}\hat{\mu}_{t+k} + (1-\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \mathbb{E}_{jt}\pi_{t+k} \\
&= (1-\theta)\mathbb{E}_{jt}[\pi_t - (1-\beta\theta)\Theta\hat{\mu}_t] + \beta\theta\mathbb{E}_{jt}\left\{(1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k[\pi_{t+1+k} - (1-\beta\theta)\Theta\hat{\mu}_{t+1+k}]\right\} \\
&= (1-\theta)\mathbb{E}_{jt}[\pi_t - (1-\beta\theta)\Theta\hat{\mu}_t] + \beta\theta\mathbb{E}_{jt}\left\{(1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k\mathbb{E}_{j,t+1}[\pi_{t+1+k} - (1-\beta\theta)\Theta\hat{\mu}_{t+1+k}]\right\} \\
&= -(1-\theta)(1-\beta\theta)\Theta\mathbb{E}_{jt}\hat{\mu}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{j,t+1} \qquad \text{(F.16)}
\end{aligned}$$

where $\pi_t = \int_{\mathcal{I}_f} \pi_{jt}\, dj$.

Note that I can write the deviation between average and desired markups as

$$\begin{aligned}
\mu_t &= p_t - \psi_t \\
&= p_t - w_t + w_t - \psi_t \\
&= -(w_t - p_t) + w_t - [w_t - a_t + \alpha n_t - \log(1-\alpha)] \\
&= -(\sigma y_t + \varphi n_t) + [a_t - \alpha n_t + \log(1-\alpha)] \\
&= -\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right) y_t + \frac{1+\varphi}{1-\alpha} a_t + \log(1-\alpha)
\end{aligned}$$

As in the benchmark model, under flexible prices ($\theta = 0$) the average markup is constant and equal to the desired μ. Consider the natural level of output, y^n_t as the equilibrium level under flexible prices and full-information rational expectations. Rewriting the above condition under the natural equilibrium,

$$\mu = -\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right) y^n_t + \frac{1+\varphi}{1-\alpha} a_t + \log(1-\alpha)$$

which I can write as

$$y_t^n = \psi_{ya} a_t + \psi_y$$

where $\psi_{ya} = \frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha}$ and $\psi_y = -\frac{(1-\alpha)[\mu-\log(1-\alpha)]}{\sigma(1-\alpha)+\varphi+\alpha}$. Therefore, I can write

$$\hat{\mu}_t = -\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)\tilde{y}_t$$

where $\tilde{y}_t = y_t - y_t^n$ is defined as the output gap. Finally, I can write the individual Phillips curve as

$$\begin{aligned}\pi_{jt} &= (1-\theta)(1-\beta\theta)\Theta\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)\mathbb{E}_{jt}\tilde{y}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{i,t+1} \\ &= \kappa\theta\mathbb{E}_{jt}\tilde{y}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{i,t+1} \end{aligned} \tag{F.17}$$

where $\kappa = \frac{(1-\theta)(1-\beta\theta)}{\theta}\Theta\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)$, and the aggregate Phillips curve can be written as

$$\pi_t = \kappa\theta\sum_{k=0}^{\infty}(\beta\theta)^k\overline{\mathbb{E}}_t^f\tilde{y}_{t+k} + (1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k\overline{\mathbb{E}}_t^f\pi_{t+k} \tag{F.18}$$

Individual and Aggregate DIS curve In order to derive the DIS curve, let us first log-linearize the profit of the monopolist. The profit D_{jt} of monopolist j at time t is

$$\begin{aligned} D_{jt} &= \frac{1}{P_t}\left(P_{jt}Y_{jt} - W_t N_{jt}\right) \\ &= \frac{P_{jt}}{P_t}Y_{jt} - W_t^r N_{jt} \end{aligned}$$

Log-linearizing around a zero-inflation steady state

$$D_j d_{jt} = \frac{P_j}{P}Y_j(p_{jt} + y_{jt} - p_t) - \frac{W^r}{P}N_j(w_t^r + n_{jt})$$

Aggregating the above expression across firms

$$\begin{aligned} y_t &= \frac{W^r N}{Y}(w_t^r + n_t) + \frac{D}{Y}d_t \\ &= \Omega(w_t^r + n_t) + (1-\Omega)d_t \end{aligned} \tag{F.19}$$

Aggregating the labor supply condition (F.6) across households, and using the goods market clearing condition

$$w_t^r = \sigma y_t + \varphi n_t$$

Inserting the above condition in (F.19), I can write

$$y_t = \frac{\Omega(1+\varphi)}{\varphi + \Omega\sigma} w_t^r + \frac{(1-\Omega)\varphi}{\varphi + \Omega\sigma} d_t$$

Introducing this last expression into the aggregate consumption function (F.9), using again the goods market clearing condition

$$y_t = -\frac{\beta}{\sigma} \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h r_{t+k} + (1-\beta) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h y_{t+k} \tag{F.20}$$

Let us now derive the DIS curve. Substracting the natural level of output from (F.20), I obtain

$$\tilde{y}_t = -\frac{\beta}{\sigma} \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h (r_{t+k} - r_{t+k}^n) + (1-\beta) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h \tilde{y}_{t+k} \tag{F.21}$$

I now need to derive an expression for the natural real interest rate. Recall that in a natural equilibrium with no price nor information frictions, the natural real interest rate is given by

$$\begin{aligned} r_t^n &= \sigma \mathbb{E}_t \Delta y_{t+1}^n \\ &= \sigma \psi_{ya} \mathbb{E}_t \Delta a_{t+1} \\ &= \sigma \psi_{ya} (\rho_a - 1) a_t \end{aligned} \tag{F.22}$$

Finally, the aggregate DIS curve is given by

$$\tilde{y}_t = -\frac{\beta}{\sigma} \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h (i_{t+k} - \pi_{t+k+1}) + (1-\beta) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h \tilde{y}_{t+k} - \psi_{ya}(1-\rho_a) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^h a_{t+k} \tag{F.23}$$

Notice that in this case there is no direct individual DIS curve. However, one can show

that the following consumption function

$$c_{it} = -\frac{\beta}{\sigma}\mathbb{E}_{it}r_t + (1-\beta)\mathbb{E}_{it}c_t + \beta\mathbb{E}_{it}c_{i,t+1} - \psi_{ya}(1-\rho_a)\mathbb{E}_{it}a_t, \quad \text{with } c_t = \int c_{it}\, di \tag{F.24}$$

is equivalent to (F.23) provided that $\lim_{T\to\infty}\beta^T\mathbb{E}_{it}c_{i,t+T}$, which is broadly assumed in the literature given $\beta < 1$.

Monetary Authority The model is closed through a Central Bank reaction function. Following Taylor (1993, 1999) I model the reaction function in terms of elasticities. The Central Bank reacts to excess inflation and output gap through a set of parameters $\{\phi_\pi, \phi_y\}$. On top of that, the monetary authority controls an exogenous component, v_t, which I model in reduced-form as an AR(1) process to account for interest rate inertia and depends on monetary shocks $\varepsilon_t^v \sim \mathcal{N}(0, \sigma_v^2)$ that are serially uncorrelated. Formally, I can write the Taylor rule as (3.7)-(3.8).

F.1.4 Discussion on Model Derivation and FIRE

Notice that throughout the model derivation I have not discussed how are beliefs and expectations formed. Therefore, the model derived above, consisting of equations (F.23), (F.18), (3.7), (3.8) and (F.11), should be interpreted as a general framework.

Under the assumption that expectations satisfy the Law of Iterated expectations, $\mathbb{E}_t[\mathbb{E}_{t+k}(\cdot)] = \mathbb{E}_t(\cdot)$ for $k > 0$, and that they are common across agents, $\overline{\mathbb{E}}_t^h(\cdot) = \overline{\mathbb{E}}_t^f(\cdot) = \mathbb{E}_t(\cdot)$, I can write the model in its usual form

$$\begin{aligned}
\tilde{y}_t &= -\frac{1}{\sigma}(i_t - \mathbb{E}_t\pi_{t+1}) + \mathbb{E}_t\tilde{y}_{t+1} + \psi_{ya}(\rho_a - 1)a_t \\
\pi_t &= \kappa\tilde{y}_t + \beta\mathbb{E}_t\pi_{t+1}
\end{aligned}$$

together with (3.7), (3.8) and (F.11).

F.2 The (FIRE) Trend-Inflation New Keynesian Model

F.2.1 Households

There is a continuum of infinitely-lived, ex-ante identical households indexed by $i \in \mathcal{I}_h = [0, 1]$ seeking to maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(C_{it}, N_{it}) \tag{F.25}$$

where utility takes a standard CRRA shape $U(C,N) = \frac{C^{1-\sigma}}{1-\sigma} - \frac{N^{1+\varphi}}{1+\varphi}$. Furthermore, the consumption index C_{it} is given by

$$C_{it} = \left(\int_{\mathcal{I}_f} C_{ijt}^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}}$$

with C_{ijt} denoting the quantity of good j consumed by household i in period t, and ϵ denotes the elasticity between goods. Here I have assumed that each consumption good is indexed by $j \in \mathcal{I}_f = [0,1]$. Given the different good varieties, the household must decide how to optimally allocate its limited expenditure on each good j. A cost-minimization problem yields

$$C_{ijt} = \left(\frac{P_{jt}}{P_t} \right)^{-\epsilon} C_{it} \tag{F.26}$$

where the aggregate price index is defined as $P_t \equiv \left(\int_{\mathcal{I}_f} P_{jt}^{1-\epsilon} dj \right)^{\frac{1}{1-\epsilon}}$. Using the above conditions, one can show that

$$\int_{\mathcal{I}_f} P_{jt} C_{ijt} \, dj = P_t C_{it}$$

I can now state the household-level budget constraint. In real terms, households decide how much to consume, work and save subject to the following restriction

$$C_{it} + B_{it} = \frac{I_{t-1}}{\Pi_t} B_{i,t-1} + \boldsymbol{w}_{it}^r N_{it} + D_{it} \tag{F.27}$$

where N_{it} denotes employment (or hours worked) by household i, B_{it} denotes savings (or bond purchases) by household i, I_{t-1} denotes the gross nominal return on savings, $\Pi_t \equiv P_t/P_{t-1}$ denotes gross inflation rate at time t, W_{it}^r denotes the realwage received by household i at time t, and D_{it} denotes dividends received by household i from the profits produced by firms. In order to avoid a potential Grossman-Stiglitz paradox, I follow the literature and noise up individual wages and dividends, so that agents cannot infer aggregate wages and output from their individual measure. Formally, I assume that wages and dividends have an aggregate and an iid idiosyncratic component, such that $X_{it} = X_t \zeta_{it}$. The optimality conditions from the household problem satisfy

$$C_{it}^{-\sigma} = \beta \mathbb{E}_t \left(\frac{R_t}{\Pi_{t+1}} C_{i,t+1}^{-\sigma} \right)$$
$$N_{it}^{\varphi} = \boldsymbol{w}_{it}^{r} C_{it}^{-\sigma}$$

Aggregating across households and log-linearizing the above conditions around a steady state with trend inflation I find

$$y_t = \mathbb{E}_t y_{t+1} - \frac{1}{\sigma}(i_t - \mathbb{E}_t \pi_{t+1})$$
$$w_t^r = \varphi n_t + \sigma y_t \tag{F.28}$$

where $x_t = \log X_t - \log X$.

F.2.2 Firms

As in the household sector, I assume a continuum of firms indexed by $j \in \mathcal{I}_f = [0, 1]$. Each firm is a monopolist producing a differentiated intermediate-good variety, producing output Y_{jt} and setting nominal price P_{jt} and making real profit D_{jt}. Technology is represented by the production function

$$Y_{jt} = A_t N_{jt}^{1-\alpha} \tag{F.29}$$

where A_t is the level of technology, common to all firms, which evolves according to

$$a_t = \rho_a a_{t-1} + \varepsilon_t^a \tag{F.30}$$

where $\varepsilon_t^a \sim \mathcal{N}(0, \sigma_a^2)$.

Aggregate Price Dynamics As in the benchmark NK model, price rigidities take the form of Calvo-lottery friction. At every period, each firm is able to reset their price with probability $(1 - \theta)$, independent of the time of the last price change. However, a firm that is unable to re-optimize gets to reset its price to a partial indexation on past inflation. Formally,

$$P_{jt} = P_{j,t-1} \Pi_{t-1}^{\omega}$$

where ω is the elasticity of prices with respect to past inflation. As a result, a firm that last reset its price in period t will face a nominal price in period $t+k$ of $P_t^* \chi_{t,t+k}$, where

$$\chi_{t,t+k} = \begin{cases} \Pi_t^{\omega} \Pi_{t+1}^{\omega} \Pi_{t+2}^{\omega} \cdots \Pi_{t+k-1}^{\omega} & \text{if } k \geq 1 \\ 1 & \text{if } k = 0 \end{cases}$$

Such environment implies that aggregate price dynamics are given by

$$P_t = \left[\theta \Pi_{t-1}^{(1-\epsilon)\omega} P_{t-1}^{1-\varepsilon} + (1-\theta)(P_{jt}^*)^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}}$$

Dividing by P_t and rearranging terms, I can write

$$\frac{P_{jt}}{P_t} = \left[\frac{1 - \theta \Pi_{t-1}^{(1-\epsilon)\omega} \Pi_t^{\epsilon-1}}{1-\theta} \right]^{\frac{1}{1-\epsilon}}$$

Log-linearizing the above expression around a steady-state with trend inflation I obtain

$$p_{jt}^* - p_t = \frac{\theta \bar{\pi}^{(\epsilon-1)(1-\omega)}}{1 - \theta \bar{\pi}^{(\epsilon-1)(1-\omega)}} \left(\pi_t - \omega \pi_{t-1} \right) \tag{F.31}$$

Optimal Price Setting A firm re-optimizing in period t will choose the price P_{jt}^* that maximizes the current market value of the profits generated while the price remains effective. Formally,

$$P_{jt}^* = \arg\max_{P_{jt}} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_{jt} \left\{ \Lambda_{t,t+k} \frac{1}{P_{t+k}} \left[P_{jt} \chi_{t,t+k} Y_{j,t+k|t} - \mathcal{C}_{t+k}(Y_{j,t+j|t}) \right] \right\}$$

subject to the sequence of demand schedules

$$Y_{j,t+k|t} = \left(\frac{P_{jt} \chi_{t,t+k}}{P_{t+k}} \right)^{-\epsilon} C_{t+k}$$

where $\Lambda_{t,t+k} \equiv \beta^k \left(\frac{C_{t+k}}{C_t} \right)^{-\sigma}$ is the stochastic discount factor, $\mathcal{C}_t(\cdot)$ is the (nominal) cost function where

$$\mathcal{C}_{t+k} = W_{t+k} N_{j,t+k|t}$$

$$= W_{t+k}\left(\frac{Y_{j,t+k|t}}{A_{t+k}}\right)^{\frac{1}{1-\alpha}}$$

and $Y_{j,t+k|t}$ denotes output in period $t+k$ for a firm j that last reset its price in period t. The First-Order Condition is

$$\sum_{k=0}^{\infty}\theta^k\mathbb{E}_{jt}\left\{\Lambda_{t,t+k}\left[(1-\epsilon)(P^*_{jt})^{-\epsilon}\left(\frac{\chi_{t,t+k}}{P_{t+k}}\right)^{1-\epsilon}Y_{j,t+k|t}+\frac{\epsilon}{1-\alpha}(P^*_{jt})^{\frac{\alpha-1-\epsilon}{1-\alpha}}\frac{W_{t+k}}{P_{t+k}}\left(\frac{Y_{j,t+k|t}}{A_{t+k}}\right)^{\frac{1}{1-\alpha}}\left(\frac{\chi_{t,t+k}}{P_{t+k}}\right)^{-\frac{\epsilon}{1-\alpha}}\right]\right\}=0$$

where $\Psi_{j,t+k|t}\equiv\mathcal{C}'_{t+k}(Y_{j,t+j|t})$ denotes the (nominal) marginal cost for firm j,

$$\Psi_{j,t+k|t}=\frac{1}{1-\alpha}A_{t+k}^{-\frac{1}{1-\alpha}}W_{t+k}Y_{j,t+k|t}^{\frac{\alpha}{1-\alpha}}$$

The FOC can be rewritten as

$$(P^*_{it})^{\frac{1-\alpha+\epsilon\alpha}{1-\alpha}}=\mathcal{M}\frac{1}{1-\alpha}\frac{\mathbb{E}_t\sum_{k=0}^{\infty}\theta^k\Lambda_{t,t+k}\frac{W_{t+k}}{P_{t+k}}\left(\frac{Y_{j,t+k|t}}{A_{t+k}}\right)^{\frac{1}{1-\alpha}}\left(\frac{\chi_{t,t+k}}{P_{t+k}}\right)^{-\frac{\epsilon}{1-\alpha}}}{\mathbb{E}_t\sum_{k=0}^{\infty}\theta^k\Lambda_{t,t+k}\left(\frac{\chi_{t,t+k}}{P_{t+k}}\right)^{1-\epsilon}Y_{j,t+k|t}}$$

where $\mathcal{M}=\frac{\epsilon}{\epsilon-1}$. Diving the above expression by $P_t^{\frac{1-\alpha+\epsilon\alpha}{1-\alpha}}=P_t^{1-\epsilon+\frac{\epsilon}{1-\alpha}}=P_t^{1-\epsilon}P_t^{\frac{\epsilon}{1-\alpha}}$,

$$\begin{aligned}\left(\frac{P^*_{it}}{P_t}\right)^{\frac{1-\alpha+\epsilon\alpha}{1-\alpha}}&=\mathcal{M}\frac{1}{1-\alpha}\frac{\mathbb{E}_t\sum_{k=0}^{\infty}\theta^k\Lambda_{t,t+k}\frac{W_{t+k}}{P_{t+k}}\left(\frac{Y_{j,t+k|t}}{A_{t+k}}\right)^{\frac{1}{1-\alpha}}\left(\chi_{t,t+k}^{-\frac{1-\omega}{\omega}}\Pi_t\right)^{-\frac{\epsilon}{1-\alpha}}}{\mathbb{E}_t\sum_{k=0}^{\infty}\theta^k\Lambda_{t,t+k}\left(\chi_{t,t+k}^{-\frac{1-\omega}{\omega}}\Pi_t\right)^{1-\epsilon}Y_{j,t+k|t}}\\&=\frac{\mathcal{M}}{1-\alpha}\frac{\Psi_t}{\Phi_t}\end{aligned}\tag{F.32}$$

where the auxiliary variables are defined, recursively, as

$$\begin{aligned}\Psi_t&\equiv\mathbb{E}_t\sum_{j=0}^{\infty}(\beta\theta)^kY_{j,t+k|t}^{\frac{1-\sigma(1-\alpha)}{1-\alpha}}A_{t+k}^{-\frac{1}{1-\alpha}}\frac{W_{t+k}}{P_{t+k}}\left(\chi_{t,t+k}^{-\frac{1-\omega}{\omega}}\Pi_t\right)^{-\frac{\epsilon}{1-\alpha}}\\&=\frac{W_t}{P_t}A_t^{-\frac{1}{1-\alpha}}Y_{jt|t}^{\frac{1-\sigma(1-\alpha)}{1-\alpha}}+\beta\theta\Pi_t^{-\frac{\epsilon\omega}{1-\alpha}}\mathbb{E}_t\left[\Pi_{t+1}^{\frac{\epsilon}{1-\alpha}}\Psi_{t+1}\right]\end{aligned}\tag{F.33}$$

$$\begin{aligned}\Phi_t&\equiv\mathbb{E}_t\sum_{j=0}^{\infty}(\beta\theta)^kY_{j,t+k|t}^{1-\sigma}\left(\chi_{t,t+k}^{-\frac{1-\omega}{\omega}}\Pi_t\right)^{1-\epsilon}\\&=Y_{jt|t}^{1-\sigma}+\beta\theta\Pi_t^{\omega(1-\epsilon)}\mathbb{E}_t\left[\Pi_{t+1}^{\epsilon-1}\Phi_{t+1}\right]\end{aligned}\tag{F.34}$$

epsilon Log-linearizing (F.32), (F.33) and around a steady state with trend inflation yields, respectively

$$\psi_t - \phi_t = \frac{1-\alpha+\epsilon\alpha}{1-\alpha}(p^*_{jt} - p_t) \tag{F.35}$$

$$\begin{aligned}\psi_t = &\left[1-\theta\beta\bar{\pi}^{\frac{\epsilon(1-\rho)}{1-\alpha}}\right]\left(w^r_t - \frac{1}{1-\alpha}a_t + \frac{1-\sigma(1-\alpha)}{1-\alpha}y_t\right)\\ &+\theta\beta\bar{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\left(\mathbb{E}_t\psi_{t+1} + \frac{\epsilon}{1-\alpha}\mathbb{E}_t\pi_{t+1} - \frac{\omega\epsilon}{1-\alpha}\pi_t\right)\end{aligned} \tag{F.36}$$

$$\phi_t = \left[1-\theta\beta\bar{\pi}^{(\epsilon-1)(1-\omega)}\right](1-\sigma)y_t + \theta\beta\bar{\pi}^{(\epsilon-1)(1-\omega)}\left[\omega(1-\epsilon)\pi_t + \mathbb{E}_t\phi_{t+1} + (\epsilon-1)\mathbb{E}_t\pi_{t+1}\right] \tag{F.37}$$

F.2.3 Equilibrium

Market clearing in the goods market implies that $Y_{jt} = C_{jt} = \int_{\mathcal{I}_h} C_{ijt}\, di$ for each j good/firm. Aggregating across firms, I obtain the aggregate market clearing condition: since assets are in zero net supply and there is no capital, investment, government consumption nor net exports, production equals consumption:

$$\int_{\mathcal{I}_f} Y_{jt}\, dj = \int_{\mathcal{I}_h}\int_{\mathcal{I}_f} C_{ijt}\, dj\, di \implies Y_t = C_t$$

Aggregate employment is given by the sum of employment across firms, and must meet aggregate labor supply

$$N_t = \int_{\mathcal{I}_h} N_{it}\, di = \int_{\mathcal{I}_f} N_{jt}\, dj$$

Using the production function (F.29) and (F.26) together with goods market clearing

$$\begin{aligned}N_t &= \int_{\mathcal{I}_f}\left(\frac{Y_{jt}}{A_t}\right)^{\frac{1}{1-\alpha}} dj\\ &= \left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}}\int_{\mathcal{I}_f}\left(\frac{P_{jt}}{P_t}\right)^{-\frac{\epsilon}{1-\alpha}} dj\\ &= \left(\frac{Y_t}{A_t}\right)^{\frac{1}{1-\alpha}} S_t\end{aligned} \tag{F.38}$$

where S_t is a measure of price dispersion and is bounded below one (see Schmitt-Grohe and Uribe (2005)). Price dispersion can be understood as the resource costs coming from price dispersion: the smaller S_t, the larger labor amount is necessary to achieve a particular level

of production. In the benchmark model with no trend inflation, $\Pi = \overline{\pi} = 1$ and S_t does not affect real variables up to the first order. Schmitt-Grohe and Uribe (2005) show that relative price dispersion can be written as

$$S_t = (1-\theta)\left(\frac{P^*_{jt}}{P_t}\right)^{-\frac{\epsilon}{1-\alpha}} + \theta \Pi_{t-1}^{-\frac{\epsilon\omega}{1-\alpha}} \Pi_t^{\frac{\epsilon}{1-\alpha}} S_{t-1} \tag{F.39}$$

Log-linearizing (F.38) and (F.39) around a steady state with trend inflation I can write, respectively

$$n_t = s_t + \frac{1}{1-\alpha}(y_t - a_t) \tag{F.40}$$

$$s_t = -\frac{\epsilon}{1-\alpha}\left(1-\theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\right)(p^*_{jt} - p_t) + \theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\left(-\frac{\epsilon\omega}{1-\alpha}\pi_{t-1} + \frac{\epsilon}{1-\alpha}\pi_t + s_{t-1}\right) \tag{F.41}$$

Aggregate DIS and Phillips Curves Combining the intratemporal labor supply condition (F.28) and the production function (F.40), I can write real wages as

$$w^r_t = \varphi s_t + \frac{\varphi + \sigma(1-\alpha)}{1-\alpha} y_t - \frac{\varphi}{1-\alpha} a_t \tag{F.42}$$

Combining the optimal price setting rule (F.35) and the aggregate price dynamics condition (F.31), denoting $\Delta_t = \pi_t - \omega\pi_{t-1}$, I can write ϕ_t in terms of Δ_t,

$$\phi_t = \psi_t - \frac{1-\alpha+\epsilon\alpha}{1-\alpha}\frac{\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}}{1-\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}}\Delta_t \tag{F.43}$$

Combining the price dispersion dynamics (F.41) and the aggregate price dynamics condition (F.31), I can write current price dispersion as a backward-looking equation in inflation and price dispersion. This equation, which does not affect real variables in the benchmark model, will be key in order to generate anchoring,

$$\begin{aligned}
s_t &= -\frac{\epsilon}{1-\alpha}\left(1-\theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\right)\frac{\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}}{1-\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}}\Delta_t + \theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\left(\frac{\epsilon}{1-\alpha}\Delta_t + s_{t-1}\right) \\
&= -\frac{\epsilon}{1-\alpha}\left[\left(1-\theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\right)\frac{\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}}{1-\theta\overline{\pi}^{(\epsilon-1)(1-\omega)}} - \theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\right]\Delta_t + \theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}} s_{t-1} \\
&= \frac{\epsilon}{1-\alpha}\frac{\delta-\chi}{1-\chi}\Delta_t + \delta s_{t-1}
\end{aligned}$$

where $\delta(\overline{\pi}) = \theta\overline{\pi}^{\frac{\varepsilon(1-\omega)}{1-\alpha}}$, $\chi(\overline{\pi}) = \theta\overline{\pi}^{(\varepsilon-1)(1-\omega)}$.

Inserting the real wage equation (F.42) into the net present value of marginal costs (F.36)

$$\begin{aligned}\psi_t &= \left[1-\theta\beta\bar{\pi}^{\frac{\epsilon(1-\rho)}{1-\alpha}}\right]\left[\varphi s_t + \frac{1+\varphi}{1-\alpha}(y_t - a_t)\right] + \theta\beta\bar{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}\left(\mathbb{E}_t\psi_{t+1} + \frac{\epsilon}{1-\alpha}\mathbb{E}_t\Delta_{t+1}\right)\\ &= (1-\beta\delta)\left[\varphi s_t + \frac{1+\varphi}{1-\alpha}(y_t - a_t)\right] + \beta\delta\left(\mathbb{E}_t\psi_{t+1} + \frac{\epsilon}{1-\alpha}\mathbb{E}_t\Delta_{t+1}\right)\end{aligned}$$

Finally, introducing (F.43) into (F.37), I can write the New Keynesian Phillips curve,

$$\Delta_t = \Theta\frac{1-\chi}{\chi}\psi_t - \Theta(1-\sigma)\frac{(1-\chi)(1-\beta\chi)}{\chi}y_t - \Theta\beta(1-\chi)\mathbb{E}_t\psi_{t+1} - \left[\Theta(\epsilon-1)\beta(1-\chi) - \beta\chi\right]\mathbb{E}_t\Delta_{t+1}$$

where $\Theta = \frac{1-\alpha}{1-\alpha+\varepsilon\alpha}$.

Monetary Authority The model is closed through a Central Bank reaction function. Following Taylor (1993, 1999) I model the reaction function in terms of elasticities. The Central Bank reacts to excess inflation and output gap through a set of parameters $\{\phi_\pi, \phi_y\}$. On top of that, the monetary authority controls an exogenous component, the monetary policy shock $\varepsilon_t^v \sim \mathcal{N}(0, \sigma_v^2)$ that are serially uncorrelated. Formally, I can write the Taylor rule as

$$i_t = \rho_i i_{t-1} + (1-\rho_i)(\phi_\pi\pi_t + \phi_y y_t) + \varepsilon_t^v \tag{F.44}$$

Steady State In steady-state the model exhibits trend inflation. The model consists of 5 equations and 5 variables, which can be written in steady-state as

$$\begin{aligned}Y &= \left[\frac{(\epsilon-1)(1-\alpha)A^{\frac{1+\varphi}{1-\alpha}}}{\epsilon S^\varphi}\right]^{\frac{1-\alpha}{\varphi+\sigma+\alpha(1-\sigma)}} = \left[\frac{(\epsilon-1)(1-\alpha)}{\epsilon S^\varphi}\right]^{\frac{1-\alpha}{\varphi+\sigma+\alpha(1-\sigma)}}\\ \Pi &= \bar{\pi}\\ 1+i &= \frac{\bar{\pi}}{\beta}\\ \Psi &= \frac{S^\varphi A^{-\frac{1+\varphi}{1-\alpha}}Y^{\frac{1+\varphi}{1-\alpha}}}{1-\theta\beta\bar{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}} = \frac{S^\varphi Y^{\frac{1+\varphi}{1-\alpha}}}{1-\theta\beta\bar{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}} = \frac{S^\varphi Y^{\frac{1+\varphi}{1-\alpha}}}{1-\beta\delta}\\ S &= \frac{1-\theta}{1-\theta\bar{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}}\left[\frac{1-\theta\bar{\pi}^{(\epsilon-1)(1-\omega)}}{1-\theta}\right]^{\frac{\epsilon}{(\epsilon-1)(1-\alpha)}} = \frac{1-\theta}{1-\delta}\left(\frac{1-\chi}{1-\theta}\right)^{\frac{\epsilon}{(\epsilon-1)(1-\alpha)}}\end{aligned}$$

hence, I can write

$$
\begin{aligned}
y &= \frac{1-\alpha}{\varphi+\sigma+\alpha(1-\sigma)}\left[\log\frac{(\epsilon-1)(1-\alpha)}{\epsilon}-\varphi s\right] \\
\pi &= \log\bar{\pi} \\
i &= \log\bar{\pi}-\log\beta=\pi-\log\beta \\
\psi &= \frac{1+\varphi}{1-\alpha}y+\varphi s-\log(1-\beta\delta) \\
s &= \log\frac{1-\theta}{1-\delta}+\frac{\epsilon}{(\epsilon-1)(1-\alpha)}\log\frac{1-\chi}{1-\theta}
\end{aligned}
$$

G Extensions to the Benchmark New Keynesian Model

G.1 Forward-Looking Models

G.1.1 Benchmark New Keynesian Model

Inserting the Taylor rule (3.7) into the DIS curve (3.6), one can write the model as a system of two first-order stochastic difference equations,

$$\boldsymbol{x}_t = \boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1}+\boldsymbol{\varphi}v_t \tag{G.1}$$

where $\boldsymbol{x}=[y_t \quad \pi_t]'$ is a 2×1 vector containing output and inflation, $\boldsymbol{\delta}$ is a 2×2 coefficient matrix and $\boldsymbol{\varphi}$ is a 2×1 vector satisfying

$$\boldsymbol{\delta}=\frac{1}{\sigma+\phi_y+\kappa\phi_\pi}\begin{bmatrix}\sigma & 1-\beta\phi_\pi \\ \sigma\kappa & \kappa+\beta(\sigma+\phi_y)\end{bmatrix}, \qquad \boldsymbol{\varphi}=\frac{1}{\sigma+\phi_y+\kappa\phi_\pi}\begin{bmatrix}1\\ \kappa\end{bmatrix}$$

The system of first-order stochastic difference equations (G.1) can be solved analytically, which is of help for our purpose. In particular, the solution to the above system of equations satisfies $\boldsymbol{x}_t=\Psi v_t$, where $\Psi=[\psi_y \quad \psi_\pi]'$ with ψ_π defined in (D.2) and

$$\psi_y=-\frac{1-\rho_v\beta}{(1-\rho_v\beta)[\sigma(1-\rho)+\phi_y]+\kappa(\phi_\pi-\rho)}$$

Model Parameters Model parameters are set to their standard values in the NK literature, reported in Table OA.1.

Parameter	Description	Value	Source/Target
σ	IES	1	Gali (2015)
β	Discount factor	0.99	Gali (2015)
φ	Inverse Frisch elasticity	5	Gali (2015)
$1-\alpha$	Labor share	1/4	Gali (2015)
ϵ	CES between varieties	9	Gali (2015)
θ	Calvo lottery	0.872	$\kappa = 0.06$
ρ	Monetary shock persistence	0.5	Gali (2015)
ϕ_π	Inflation coefficient Taylor rule	1.5	Gali (2015)
ϕ_y	Output gap coefficient Taylor rule	0.5	Stability
σ_ε	Volatility monetary shock	1	Gali (2015)

Table OA.1: Model parameters.

G.1.2 Accommodating Technology and Cost-push Shocks

In this section I extend the general model to accommodate cost-push shocks. The demand side is still described by (F.23), which under the FIRE assumption collapses to

$$\tilde{y}_t = -\frac{1}{\sigma}(i_t - \mathbb{E}_t\pi_{t+1}) - (1-\rho_a)\psi_{ya}a_t + \mathbb{E}_t\tilde{y}_{t+1} \tag{G.2}$$

In order to accommodate cost-push shocks in a micro-consistent manner, I allow the elasticity of substitution among goof varieties, ϵ, to vary over time according to some stationary process $\{\epsilon_t\}$. Assuming constant returns to scale in the production function (F.10) ($\alpha = 0$) for simplicity, the Phillips curve becomes

$$\begin{aligned}\pi_t &= \beta\mathbb{E}_t\pi_{t+1} - \lambda\hat{\mu}_t + \lambda\hat{\mu}_t^n \\ &= \beta\mathbb{E}_t\pi_{t+1} + \kappa\tilde{y}_t + u_t\end{aligned} \tag{G.3}$$

where $\mu_t^n = \log\frac{\epsilon_t}{\epsilon_t - 1}$ is the time-varying desired markup and $\hat{\mu}_t^n = \mu_t^n - \mu$. I assume that the exogenous process $u_t = \lambda\hat{\mu}_t^n$ follows an AR(1) process with autorregressive coefficient ρ_u. Combining (G.2), (G.3), (3.7) and the respective shock processes, I can write the equilibrium conditions as a system of stochastic difference equations

$$\tilde{A}\boldsymbol{x}_t = \tilde{B}\mathbb{E}_t\boldsymbol{x}_{t+1} + \tilde{C}\boldsymbol{w}_t \tag{G.4}$$

where $\boldsymbol{x}_t = [y_t \quad \pi_t]'$ is a 2×1 vector containing output and inflation, $\boldsymbol{w}_t = [v_t \quad a_t \quad u_t]'$ is a 3×1 vector containing the monetary, technology and cost-push shocks, $\tilde{A}$ is a 2×2

coefficient matrix, $\tilde{B}$ is a 2×2 coefficient matrix and $\tilde{C}$ is a 2×3 matrix satisfying

$$\tilde{A} = \begin{bmatrix} \sigma + \phi_y & \phi_\pi \\ -\kappa & 1 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \sigma & 1 \\ 0 & \beta \end{bmatrix}, \quad \text{and} \quad \tilde{C} = \begin{bmatrix} -1 & -\sigma(1-\rho_a)\psi_{ya} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Premultiplying the system by $\tilde{A}^{-1}$ I obtain

$$\boldsymbol{x}_t = \boldsymbol{\delta}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\varphi}\boldsymbol{w}_t \tag{G.5}$$

where $\boldsymbol{\delta} = \tilde{A}^{-1}\tilde{B}$ and $\boldsymbol{\varphi} = \tilde{A}^{-1}\tilde{C}$. Notice that $\boldsymbol{w}_t$ follows a VAR(1) process with autorregressive coefficient matrix $R = \text{diag}(\rho_v, \rho_a, \rho_u)$. Using the method for undetermined coefficients, the solution to (G.5) is conjectured to be of the form

$$\begin{aligned} \tilde{y}_t &= \Phi_y \boldsymbol{w}_t, \quad \text{where} \quad \Phi_y = [\phi_{yv} \quad \phi_{ya} \quad \phi_{yu}] \\ \tilde{\pi}_t &= \Phi_\pi \boldsymbol{w}_t, \quad \text{where} \quad \Phi_\pi = [\phi_{\pi v} \quad \phi_{\pi a} \quad \phi_{\pi u}] \end{aligned}$$

Imposing the conjectured relations into (G.5) allows one to solve for the undetermined coefficients ϕ_{yv}, ϕ_{ya}, ϕ_{yu}, $\phi_{\pi v}$, $\phi_{\pi a}$ and $\phi_{\pi u}$, which satisfy the following condition

$$\Phi = \boldsymbol{\delta}\Phi R + \boldsymbol{\varphi}$$

where $\Phi = [\Phi_y \quad \Phi_\pi]'$ is a 2×3 vector containing all the unknown parameters. The solution to the above system of unknown parameters satisfied

$$\begin{aligned} \phi_{yv} &= -\frac{1-\rho_v\beta}{(1-\rho_v\beta)[\sigma(1-\rho_v)+\phi_y]+\kappa(\phi_\pi-\rho_v)} \\ \phi_{ya} &= -\frac{\sigma\psi_{ya}(1-\rho_a)(1-\rho_a\beta)}{(1-\rho_a\beta)[\sigma(1-\rho_a)+\phi_y]+\kappa(\phi_\pi-\rho_a)} \\ \phi_{yu} &= -\frac{\phi_\pi-\rho_u}{(1-\rho_u\beta)[\sigma(1-\rho_u)+\phi_y]+\kappa(\phi_\pi-\rho_u)} \\ \phi_{\pi v} &= -\frac{\kappa}{(1-\rho_v\beta)[\sigma(1-\rho_v)+\phi_y]+\kappa(\phi_\pi-\rho_v)} \\ \phi_{\pi a} &= -\frac{\kappa\sigma\psi_{ya}(1-\rho_a)}{(1-\rho_a\beta)[\sigma(1-\rho_a)+\phi_y]+\kappa(\phi_\pi-\rho_a)} \\ \phi_{\pi u} &= \frac{\sigma(1-\rho_u)+\phi_y}{(1-\rho_u\beta)[\sigma(1-\rho_u)+\phi_y]+\kappa(\phi_\pi-\rho_u)} \end{aligned}$$

Parameter	Description	Value	Source/Target
ρ_a	Technology shock persistence	0.9	Galí (2015)
ρ_u	Cost-push shock persistence	0.8	Galí (2015)
$\sigma_{\varepsilon a}$	Technology innovation pre-1985	1	Galí (2015)
$\sigma_{\varepsilon u}$	Cost-push innovation	1	Galí (2015)

Table OA.2: Estimated parameters

and therefore equilibrium dynamics are given by

$$\tilde{y}_t = \phi_{yv} v_t + \phi_{ya} a_t + \phi_{yu} u_t \tag{G.6}$$

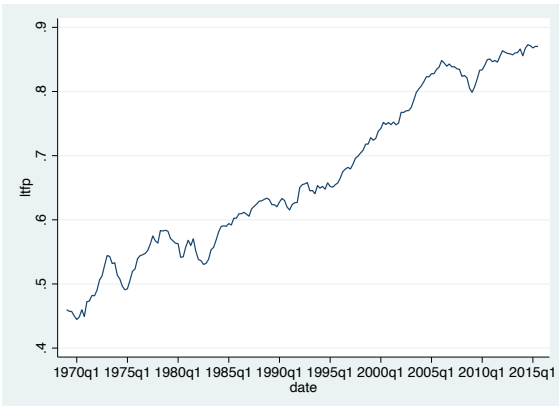
$$\pi_t = \phi_{\pi v} v_t + \phi_{\pi a} a_t + \phi_{\pi u} u_t \tag{G.7}$$

In this framework with multiple shocks, I study inflation persistence as the first-order autocorrelation coefficient ρ_1 as

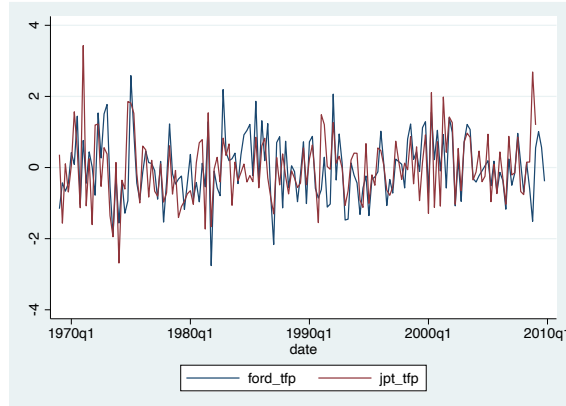
$$\rho_1 = \frac{\rho_v \frac{\phi_{\pi v}^2 \sigma_{\varepsilon v}^2}{1-\rho_v^2} + \rho_a \frac{\phi_{\pi a}^2 \sigma_{\varepsilon a}^2}{1-\rho_a^2} + \rho_u \frac{\phi_{\pi u}^2 \sigma_{\varepsilon u}^2}{1-\rho_u^2}}{\frac{\phi_{\pi v}^2 \sigma_{\varepsilon v}^2}{1-\rho_v^2} + \frac{\phi_{\pi a}^2 \sigma_{\varepsilon a}^2}{1-\rho_a^2} + \frac{\phi_{\pi u}^2 \sigma_{\varepsilon u}^2}{1-\rho_u^2}}$$

Notice that the Taylor rule coefficient ϕ_π is now relevant for the first-order autocorrelation of the inflation process. As a result, a fall in inflation persistence could be explained by a contemporaneous fall in the monetary stance ϕ_π. Having shown in Section D.1 that there is no evidence for a change in monetary policy shock persistence, I investigate below if the data suggests a structural break in technology and/or cost-push shock persistence. As I show below, I find no evidence of a fall in productivity or cost-push shocks' persistence over time. Using a textbook parameterization, reported in Table OA.1 (first six rows) and Table OA.2, I find the NK model extended with technology and cost-push shocks predicts that aggregate persistence should have decreased only moderately after an increase in ϕ_π: $\rho_{1,\text{pre}} = 0.80$ vs. $\rho_{1,\text{post}} = 0.80$. I therefore conclude that the standard model cannot explain the fall in inflation persistence after 1985.

Technology Shocks In this section I rely on the vast literature on technology shocks, dating back to Solow (1957) and Kydland and Prescott (1982). Early work in the literature generally assumed that a regression on the (log) production function reports residuals that can be interpreted as (log) TFP neutral shocks, as the one discussed in this section. Due to endogeneity concerns between capital and TFP, the literature moved forward and estimated



(a) (Log) TFP process from Fernald (2014)



(b) Technology shocks from Francis et al. (2014) and Justiniano et al. (2011)

Figure OA.1: TFP dynamics

TFP shocks through different assumptions and methods. In this new wave, Galí (1999) used long-run restrictions to identify neutral technology shocks by assuming that technology shocks are the only that can have permanent effects on labor productivity. Following this idea, Francis et al. (2014) identify technology shocks as the shock that maximizes the forecast error variance share of labor productivity at some horizon. Basu et al. (2006) instead estimate TFP by adjusting the annual Solow residual for utilization (using hours per worker as a proxy), and Fernald (2014) extended the series to quarterly frequency. Finally, Justiniano et al. (2011) obtain technology shocks by estimating a NK model, incorporating other technology-related shocks such as investment-specific technology and marginal efficiency of investment shocks. Ramey (2016) compares the shocks, and shows that the IRFs of standard aggregate variables after the each shock series are similar. In particular, Francis et al. (2014) and Justiniano et al. (2011) produce remarkably similar IRFs of real GDP, hours and consumption.

I plot the different series in Figure OA.1. Notice the difference between the left and right panels: while Fernald (2014) estimates directly (log) technology a_t, Francis et al. (2014) and Justiniano et al. (2011) estimate the technology shock ε_t^a. I overcome the difficulty with the estimation of technology persistence by estimating persistence in the natural real interest rate process. In the standard NK model, the natural real rate is given by (F.22), which can be rewritten using the AR(1) properties of the technology process as

$$r_t^n = \rho_a r_{t-1}^n - \sigma \psi_{ya}(1-\rho_a)\varepsilon_t^a \tag{G.8}$$

	(1) Technology	(2) SB	(3) Natural rate	(4) SB	(5) Natural rate	(6) SB
(Log) TFP_{t-1}	0.998^{***}	0.990^{***}				
	(0.00454)	(0.00860)				
(Log) TFP_{t-1} change		0.00323				
		(0.00339)				
Natural rate_{t-1}			0.951^{***}	0.945^{***}	0.963^{***}	0.957^{***}
			(0.0317)	(0.0327)	(0.0367)	(0.0404)
Technology shock in Francis et al. (2014)			0.0511^{**}	0.0514^{**}		
			(0.0234)	(0.0237)		
Natural rate_{t-1} change				−0.0106		−0.00863
				(0.0129)		(0.0141)
Technology shock in Justiniano et al. (2011)					0.0191	0.0195
					(0.0278)	(0.0280)
Constant	0.00360	0.00743^{*}	0.128	0.162	0.0878	0.123
	(0.00327)	(0.00445)	(0.0968)	(0.109)	(0.114)	(0.140)
Observations	186	186	163	163	160	160

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table OA.3: Regression table

I use the Federal Reserve estimate of the natural interest rate series, produced by Holston et al. (2017), as our proxy for r_t^n. Table OA.3 reports our results. The first two columns report the (direct) estimate of the technology process (F.11) persistence and its structural break around 1985:I, while columns three to six report the estimate of the natural real rate process (G.8) using the technology series constructed by Francis et al. (2014) and Justiniano et al. (2011), respectively. Our results suggest that there is no evidence for a fall in technology persistence over time.

Cost-push Shocks In the benchmark NK model with monopolistic competition among firms, cost-push shocks are interpreted as the deviation from the desired time-varying price-cost markup, which depends on the elasticity of substitution among good varieties. Nekarda and Ramey (2010) estimate the structural time-varying price-cost markup under a richer framework than the benchmark NK model. In particular, they consider both labor and capital as inputs in the production function. They argue that measured wages are a better indicator for marginal costs than labor compensation, and provide a range of markup measures depending on the elasticity of substitution between capital and labor. As a result, they obtain markup estimates either from labor side or the capital side. Since our model does not include capital, I will rely on the labor-side estimates.

Figure OA.2 plots two different measures of the cost-push shock. In the first, the authors rely on a Cobb-Douglas production function in order to estimate the markup, while in

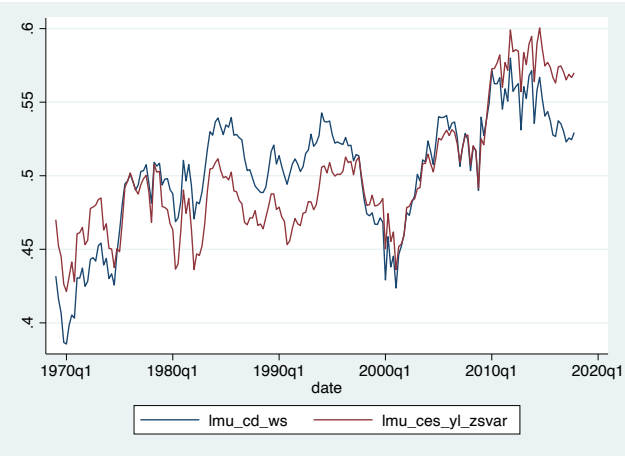


Figure OA.2: Markup series

	(1) Cobb-Douglas	(2) SB	(3) CES	(4) SB
Markup_{t-1}	0.945***	0.938***	0.963***	0.947***
	(0.0246)	(0.0305)	(0.0234)	(0.0252)
$\text{Markup}_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$		0.00187		0.00472
		(0.00436)		(0.00419)
Constant	0.0280**	0.0307**	0.0189	0.0252**
	(0.0125)	(0.0146)	(0.0117)	(0.0120)
Observations	195	195	195	195

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table OA.4: Regression table

the second the authors rely on a CES production function, estimating labor-augmented technology using long-run restrictions as in Galí (1999). I therefore estimate the first-order autocorrelation using these two measures. Our results are reported in Table OA.4. Columns one and two report the estimates based on the Cobb-Douglas production function, while columns three to four report the estimates based on the (labor-side) CES production function. I find no evidence of a change in cost-push persistence over time

G.1.3 Optimal Monetary Policy under Discretion

Following Galí (2015), the welfare losses experienced by a representative consumer, up to a second-order approximation, are proportional to

$$\mathbb{E}_0 \sum_{k=0}^{\infty} \beta^t \left(\pi_t^2 + \frac{\kappa}{\epsilon} x_t^2 \right) \tag{G.9}$$

where $x_t \equiv y_t - y_t^e$ is the welfare-relevant output gap, with $y_t^e = \psi_{ya} a_t$ denoting the (log) efficient level of output. Notice that κ/ϵ regulates the (optimal) relative weight that the social planner (or the monetary authority) assigns to the welfare-relevant output gap. In this case, the DIS can be written as

$$x_t = -\frac{1}{\sigma}(i_t - \mathbb{E}_t \pi_{t+1}) - (1 - \rho_a)\psi_{ya} a_t + \mathbb{E}_t x_{t+1} \tag{G.10}$$

I can also rewrite the Phillips curve as

$$\pi_t = \beta \mathbb{E}_t \pi_{t+1} + \kappa x_t + u_t \tag{G.11}$$

where $u_t \equiv \kappa(y_t^e - y_t^n)$. Again, I assume that the cost-push shock follows an AR(1) process with autorregressive coefficient ρ_u.

Under discretion, the central bank does not control future output gap or inflation, but just the current measures. Therefore, the monetary authority minimizes $\pi^2 + \frac{\kappa}{\epsilon} x_t^2$ subject to the constraint $\pi_t = \kappa x_t + \xi_t$, where $\xi_t \equiv \beta \mathbb{E}_t \pi_{t+1} + u_t$ is treated as a non-policy shock (one can show that $\mathbb{E}_t \pi_{t+1}$ is a function of future output gaps). The optimality condition is

$$x_t = -\epsilon \pi_t \tag{G.12}$$

In case of inflationary pressures, the Central bank will reduce output below its potential, "leaning against the wind". In this case, the welfare-relevant output gap and inflation follow

$$\tilde{y}_t = -\frac{1 - \rho_u \beta + 2\epsilon\kappa}{\kappa(1 - \rho_u \beta + \epsilon\kappa)} u_t \tag{G.13}$$

$$\pi_t = \frac{1}{1 - \rho_u \beta + \epsilon\kappa} u_t \tag{G.14}$$

Using the DIS curve (3.6) and the optimality conditions (G.13) and (G.14), I can reverse-

engineer the following Taylor rule, which replicates the optimal allocation under discretion

$$\begin{aligned} i_t &= \frac{\rho_u + \epsilon\sigma(1-\rho_u)}{1-\beta\rho_u+\epsilon\kappa} u_t - (1-\rho_a)\psi_{ya} a_t \\ &= \Psi_i u_t - (1-\rho_a)\psi_{ya} a_t \end{aligned} \tag{G.15}$$

Unfortunately, such a rule yields multiple equilibria since it does not satisfy the Taylor Principle. However, adding a component $\phi_\pi \left(\pi_t - \frac{1}{1-\rho_u\beta+\epsilon\kappa} u_t\right) = 0$, I can write

$$\begin{aligned} i_t &= \phi_\pi \pi_t + \frac{\epsilon\sigma(1-\rho_u) - (\phi_\pi - \rho_u)}{1-\beta\rho_u+\epsilon\kappa} u_t - (1-\rho_a)\psi_{ya} a_t \\ &= \phi_\pi \pi_t + \Theta_i u_t - (1-\rho_a)\psi_{ya} a_t \end{aligned} \tag{G.16}$$

Inserting condition (G.14) to eliminate the cost-push shock yields

$$\begin{aligned} i_t &= \phi_\pi \pi_t + [\epsilon\sigma(1-\rho_u) - (\phi_\pi - \rho_u)]\pi_t - (1-\rho_a)\psi_{ya} a_t \\ &= \phi_\pi \pi_t + \phi_{\pi,\mathbb{1}_{\{t\geq 1985:I\}}} \pi_t - (1-\rho_a)\psi_{ya} a_t \end{aligned} \tag{G.17}$$

As a result, one could understand the documented increase in the Taylor rule as a version of optimal discretionary policy. In our benchmark specification I find $\phi_{\pi,\mathbb{1}_{\{t\geq 1985:I\}}} = 0.95$, which aligns well with the data. I already discussed that an increase in ϕ_π does not affect inflation persistence. What if the change in the monetary stance was not a mere increase in the elasticity of nominal rates with respect to inflation, but an additional response to cost-push shocks in the Taylor rule? Recall that, under discretion, inflation dynamics are given by (G.14), which I can write as

$$\pi_t = \rho_u \pi_{t-1} + \frac{1}{1-\rho_u\beta+\epsilon\kappa} \varepsilon_t^u \tag{G.18}$$

Compared to the pre-1985 dynamics, described by (G.7) and disregarding technology shocks for simplicity, inflation persistence would be even larger if $\rho_u > \rho_v$, which I have documented in Tables A.x Panel A and OA.4. That is, optimal discretionary policy would not explain the fall in inflation persistence, provided that cost-push persistence has been stable throughout the decades, and that cost-push shocks are more persistent than monetary policy shocks, which would have generated an *increase* in inflation persistence.[61]

61. Including technology shocks in the comparison of (G.7) and (G.18) would alter the results, provided that $\rho_a > \rho_u > \rho_v$. However, since ρ_u is in between the two other highly persistent parameters and none of

G.1.4 Indeterminacy

The previous literature has considered the possibility of the Fed conducting a passive monetary policy before 1985, which in the lens of the NK framework would lead to multiplicity of equilibria. For example, Clarida et al. (2000) document that the inflation coefficient in the Taylor rule was well below one, not satisfying the Taylor principle. Lubik and Schorfheide (2004) estimate a NK model under determinacy and indeterminacy, and argue that monetary policy after 1982 is consistent with determinacy, whereas the pre-Volcker policy is not. We study here if this change in the monetary stance could have affected inflation persistence.

Consider the standard framework in (G.1). We have explored inflation dynamics under determinacy. In this section we uncover the (multiple) stable solutions under indeterminacy, where $\phi_\pi < 1 - \frac{1-\beta}{\kappa}\phi_y$. Following Lubik and Schorfheide (2003), we rewrite the model as

$$\Gamma_0 \xi_t = \Gamma_1 \xi_{t-1} + \Psi \varepsilon_t^v + \Pi \eta_t$$

where $\xi_t = [\xi_t^y \quad \xi_t^\pi \quad v_t]'$, $\eta_t = [\eta_t^y \quad \eta_t^\pi]'$ and we denote the conditional forecast $\xi_t^x = \mathbb{E}_t x_{t+1}$ and the forecast error $\eta_t^x = x_t - \xi_{t-1}^x$, with

$$\Gamma_0 = \begin{bmatrix} 1 & \frac{1}{\sigma} & -\frac{1}{\sigma} \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \Gamma_1 = \begin{bmatrix} 1+\frac{\phi_y}{\sigma} & \frac{\phi_\pi}{\sigma} & 0 \\ -\kappa & 1 & 0 \\ 0 & 0 & \rho \end{bmatrix}, \qquad \Psi = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \Pi = \begin{bmatrix} 1+\frac{\phi_y}{\sigma} & \frac{\phi_\pi}{\sigma} \\ -\kappa & 1 \\ 0 & 0 \end{bmatrix}$$

Premultiplying the system by Γ_0^{-1} we obtain the reduced-form dynamics

$$\xi_t = \Gamma_1^* \xi_{t-1} + \Psi^* \varepsilon_t^v + \Pi^* \eta_t$$

Using the Jordan decomposition of $\Gamma_1^* = J \Lambda J^{-1}$, and denoting $w_t = J^{-1}\xi_t$, we can write

$$w_t = \Lambda w_{t-1} + J^{-1}\Psi^* \varepsilon_t^v + J^{-1}\Pi^* \eta_t$$

Let the $\boldsymbol{w}_{it}$ denote ith element of w_t, $[J^{-1}\Psi^*]_i$ denote the ith row of $J^{-1}\Psi^*$ and $[J^{-1}\Pi^*]_i$ denote the ith row of $J^{-1}\Pi^*$. Since Λ is a diagonal matrix, the dynamic process can be decomposed in 3 uncoupled AR(1) processes. Define $\mathcal{I}_x$ denote the set of unstable AR(1)

them have changed over time, the difference (if any) in reduced-form persistence in (G.7) and (G.18) would be small, and would not explain the documented large fall.

processes, and let Ψ_x^J and Π_x^J be the matrices composed of the row vectors $[J^{-1}\Psi^*]_i$ and $[J^{-1}\Pi^*]_i$ such that $i \in \mathcal{I}_x$. Finally, we proceed with a singular value decomposition of the matrix Π_x^J,

$$\Pi_x^J = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} D_{11} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1' \\ V_2' \end{bmatrix} = U_1 D_1 1 V_1'$$

Lubik and Schorfheide (2003) prove that if there exists a solution in the indeterminacy region, it is of the form

$$\xi_t = \Gamma_1^* \xi_{t-1} + [\Psi^* - \Pi^* V_1 D_{11}^{-1} U_1' \Psi_x^J] \varepsilon_t^v + \Pi^* V_2 (\tilde{M} \varepsilon_t^v + M_\zeta \zeta_t)$$

Two aspects deserve a discussion. First, matrices $\tilde{M}$ and M_ζ do not depend on model parameters, which yields the multiplicity of equilibria. Following Lubik and Schorfheide (2003), we select the equilibrium that produces the same dynamics as the determinate framework on impact.[62] Second, the model features i.i.d sunspot shocks ζ_t that affect equilibrium dynamics.

In order to obtain the model dynamics, we set parameters to the values reported in Table OA.1, with the exception of ϕ_π. For the indeterminate case we set $\phi_{\pi,\text{ind}} = 0.83$, the estimate reported by Clarida et al. (2000). I find that a first-order autocorrelation coefficient of 0.643. Interestingly, the first-order autocorrelation coefficient is robust to changes in the non-fundamental response of the economy to sunspot shocks.

G.2 Backward-looking New Keynesian Models

The main reason for the failure in explaining the change in the dynamics in the benchmark NK model is that endogenous outcome variables, output gap and inflation, are proportional to the monetary policy shock process and thus inherit its dynamics. This is a result of having a pure forward-looking model. A direct consequence is that endogenous outcome variables are not intrinsically persistent, and therefore its persistence is simply inherited from the exogenous driving force. In this section I enlarge the standard NK model to accommodate a backward-looking dimension, including a lagged term $\boldsymbol{x}_{t-1}$ in the system of equations (G.1).

I do so in two different ways: in the first extension, discussed in section G.2.1, I explore a change in the monetary stance from a passive Taylor rule towards optimal policy under commitment. In the second extension, discussed in section G.2.2, I include price-indexing

62. We set $\tilde{M}$ such that $-V_1 D_{11}^{-1} U_1' \Psi_x^J + V_2 \tilde{M} = -\psi_\pi$, and M_ζ such that $V_{2,2}\zeta_0 = \psi_\pi \varepsilon_0^v$.

firms, which introduces anchoring in the supply side. In the third extension I introduce log-linearize the standard model around a steady state with trend inflation, which endogenously creates anchoring in the demand and supply sides.

G.2.1 Optimal Monetary Policy under Commitment

Our first backward-looking framework is the benchmark NK model with optimal monetary policy under commitment. Under commitment, the monetary authority can credibly control household's and firm's expectations. As a result, the Central bank program is to minimize (G.9) subject to the sequence of constraints (G.11). The optimality conditions from this program yield the following conditions relating the welfare-relevant output gap and inflation

$$x_0 = -\epsilon\pi_0 \tag{G.19}$$

$$x_t = x_{t-1} - \epsilon\pi_t \tag{G.20}$$

for $t \geq 1$. Notice that these two conditions can be jointly represented as an implicit price-level target

$$x_t = -\epsilon\hat{p}_t \tag{G.21}$$

where $\hat{p}_t \equiv p_t - p_{-1}$ is the (log) deviation of the price level from an initial target. Combining the Phillips curve (G.11) and the optimal price level target (G.21) I obtain a second-order stochastic difference equation

$$\hat{p}_t = \gamma\hat{p}_{t-1} + \gamma\beta\mathbb{E}_t\hat{p}_{t+1} + \gamma u_t$$

where $\gamma = (1 + \beta + \epsilon\kappa)^{-1}$. The stationary solution to the above condition satisfies

$$\hat{p}_t = \delta\hat{p}_{t-1} + \frac{\delta}{1 - \beta\delta\rho_u}u_t \tag{G.22}$$

where $\delta = \frac{1-\sqrt{1-4\beta\gamma^2}}{2\gamma\beta} \in (0, 1)$ is the inside root of the following lag polynomial

$$\mathcal{P}(x) = \gamma\beta x^2 - x + \gamma$$

Inserting the price level target (G.21) into (G.22), I can write the welfare-relevant output gap in terms of the cost-push shock

$$
\begin{aligned}
x_0 &= -\frac{\epsilon\delta}{1-\delta\beta\rho_u}u_0 \\
x_t &= \delta x_{t-1} - \frac{\epsilon\delta}{1-\delta\beta\rho_u}u_t
\end{aligned}
\tag{G.23}
$$

Notice that (G.25) can be written in terms of the lag polynomial as

$$
\Delta x_t = -\frac{\epsilon\delta}{1-\delta\beta\rho_u}\frac{1}{1-\delta L}\Delta u_t
$$

which I can insert back into (G.19)-(G.20) to obtain inflation dynamics

$$
\begin{aligned}
\pi_0 &= \frac{\delta}{1-\delta\beta\rho_u}u_0 \\
\pi_t &= \delta\pi_{t-1} + \frac{\delta}{1-\delta\beta\rho_u}\Delta u_t
\end{aligned}
\tag{G.24}
$$

Rewriting the output gap dynamics

$$
\tilde{y}_t = \delta\tilde{y}_{t-1} - \frac{1-\delta(\beta\rho_u - \kappa\epsilon)}{1-\delta\beta\rho_u}u_t + \frac{\delta}{\kappa}u_{t-1} \tag{G.25}
$$

Just as in the case under discretion, the monetary authority can engineer a Taylor rule that produces the optimal dynamics. Inserting (G.21), (G.22) and (G.25) into the DIS curve (G.10) I can specify the following Taylor rule,

$$
\begin{aligned}
i_t &= (1-\delta)(\sigma\epsilon - 1)\hat{p}_t - \sigma\psi_{ya}(1-\rho_a)a_t \\
&= \phi_p\hat{p}_t + \xi_t
\end{aligned}
\tag{G.26}
$$

which produces the same allocation than the optimal policy. Inserting (G.24) in the Taylor rule, I can write

$$
\begin{aligned}
i_t &= (1-\delta)(\sigma\epsilon - 1)\hat{p}_t - \sigma\psi_{ya}(1-\rho_a)a_t + \phi_\pi\left(\pi_t - \delta\pi_{t-1} - \frac{\delta}{1-\delta\beta\rho_u}\Delta u_t\right) \\
&= \phi_\pi\pi_t + (1-\delta)(\sigma\epsilon - 1)\hat{p}_t - \phi_\pi\delta\pi_{t-1} - \frac{\phi_\pi\delta}{1-\delta\beta\rho_u}\Delta u_t - \sigma\psi_{ya}(1-\rho_a)a_t \\
&= \phi_\pi\pi_t + (1-\delta)(\sigma\epsilon - 1)(\pi_t + \hat{p}_{t-1}) - \phi_\pi\delta\pi_{t-1} - \frac{\phi_\pi\delta}{1-\delta\beta\rho_u}\Delta u_t - \sigma\psi_{ya}(1-\rho_a)a_t
\end{aligned}
$$

	(1) Taylor rule	(2) SB	(3) Optimal MP (CD)	(4) Optimal MP (CES)
π_t	1.389***	1.247***	1.173***	1.169***
	(0.0659)	(0.0730)	(0.0724)	(0.0727)
$\pi_t \times \mathbb{1}_{\{t \geq t^*\}}$		0.553***	2.065**	2.018**
		(0.152)	(0.944)	(0.986)
$\pi_{t-1} \times \mathbb{1}_{\{t \geq t^*\}}$			0.581	0.598
			(0.763)	(0.752)
$p_t \times \mathbb{1}_{\{t \geq t^*\}}$			−0.00252***	−0.00243***
			(0.000794)	(0.000830)
$u_t \times \mathbb{1}_{\{t \geq t^*\}}$			−1.148*	−1.057
			(0.629)	(0.688)
Observations	203	203	192	192

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table OA.5: Regression table

$$= \phi_\pi \pi_t + \phi_{\pi,\mathbb{1}_{\{t \geq 1985:I\}}} \pi_t + \phi_{\pi,\mathbb{1}_{\{t \geq 1985:I\}}} \hat{p}_{t-1} - \phi_\pi \delta \pi_{t-1} - \frac{\phi_\pi \delta}{1 - \delta\beta\rho_u} \Delta u_t - \sigma \psi_{ya}(1 - \rho_a) a_t$$

$$= \phi_\pi \pi_t + \phi_{\pi,\mathbb{1}_{\{t \geq 1985:I\}}} \pi_t + \phi_{\pi,\mathbb{1}_{\{t \geq 1985:I\}}} \hat{p}_{t-1} - \phi_\pi \delta \pi_{t-1} - \frac{\phi_\pi \delta}{1 - \delta\beta\rho_u} \Delta u_t + \xi_t$$

where ξ_t is an AR(1) process. Our standard parameterization, reported in Table OA.1, suggests $\phi_{\pi,\mathbb{1}_{\{t \geq 1985:I\}}} = 3.56$, which is excessive considering our previous empirical findings. To confirm this, I estimate the above Taylor rule.

Table OA.5 reports our results. Columns one and two repeat our previous exercise but assuming no response to output gap deviations. Columns three to four report the estimates of the optimal Taylor rule under commitment, using Nekarda and Ramey (2010) estimates of markups. Our results support the notion that the Fed included the price level and the cost-push shock in its Taylor rule. However, the results are inconsistent with the theory, since the increase in the inflation coefficient and the increase in the price level coefficients are of opposite sign. Additionally, the change in the inflation coefficient is still far from the model-implied change that supports a commitment-rule.

Parameter	Description	Value	Source/Target
ω	Price indexation	0.75	Range literature

Table OA.6: Model Parameters

G.2.2 Price Indexation

Consider a backward-looking version of the Phillips curve, microfounded through price indexation at the firm level and governed by ω

$$\pi_t = \frac{\omega}{1+\beta\omega}\pi_{t-1} + \frac{\kappa}{1+\beta\omega}\tilde{y}_t + \frac{\beta}{1+\beta\omega}\mathbb{E}_t\pi_{t+1} \tag{G.27}$$

The rest of the model equations are the same as in the benchmark model, (3.6), (3.7) and (3.8). The model derivation is relegated to Online Appendix F.2, and the parameterization is identical to that of Table OA.1, with the model enlarged by the price-indexation parameter ω. The parameterization of such parameter is not a clear one. As I show below, price indexation implies that every price is changed every period, and therefore one could not identify the Calvo restricted firms in the data and estimate ω. As a result, the parameter is usually estimated using aggregate data and trying to match the anchoring of the inflation dynamics, and its estimate will therefore depend on the additional model equations. I set $\omega = 0.75$, which is in the range of the literature (0.21 in Smets and Wouters (2007), 1 in Christiano et al. (2005)).

The model can be collapsed to a system of three second-order stochastic difference equations

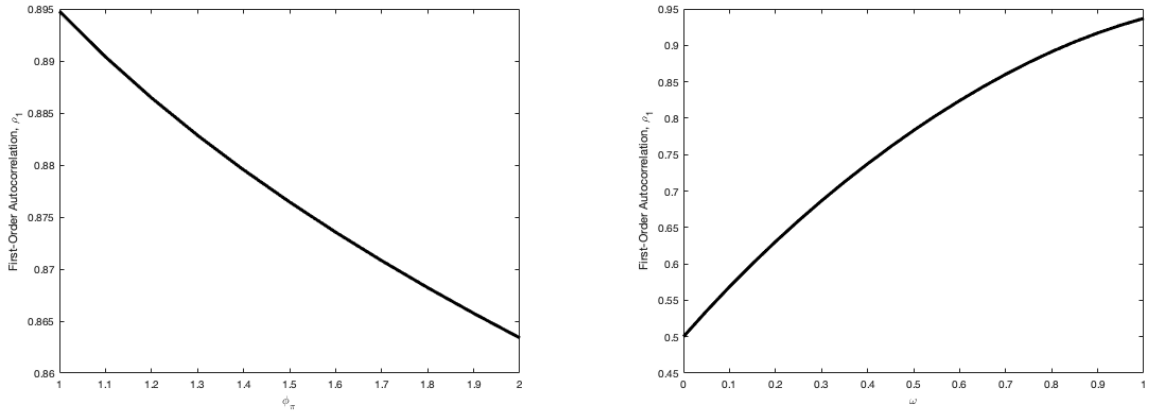
$$\boldsymbol{x}_t = \boldsymbol{\Gamma_b}\boldsymbol{x}_{t-1} + \boldsymbol{\Gamma_f}\mathbb{E}_t\boldsymbol{x}_{t+1} + \boldsymbol{\Lambda}v_t$$

where $\boldsymbol{x}_t = [y_t \quad \pi_t]'$. The solution of the above system satisfies

$$\boldsymbol{x}_t = A\boldsymbol{x}_{t-1} + \Psi v_t \tag{G.28}$$

where both matrices $A(\phi_\pi, \Phi)$ and $\Psi(\phi_\pi, \Phi)$ depend now on ϕ_π and the rest of the model parameters Φ. Notice that a key difference between the benchmark model and this backward-looking version is that a change in ϕ_π will have an effect on inflation persistence, and could therefore explain the fall in inflation persistence.

In Figure OA.3a I show that a change in the monetary policy stance has now a significant effect on inflation persistence: a change of ϕ_π from 1 to 2, as I have documented in Table



(a) Change in ϕ_π

(b) Change in ω

Figure OA.3: Inflation first-order autocorrelation in the backward-looking NK model

A.x Panel A, produces a fall in the first-order autocorrelation of inflation from around 0.895 to 0.865. However, is not enough to produce the effect that I observe in the data. The target now is to find a candidate parameter that can explain the observed loss in inflation persistence. The ideal candidate is ω, since this term produces anchoring in the Phillips curve (G.27). As I show in Figure OA.3b, as ω decreases so does inflation persistence.

I can see in Figure OA.3b that the decrease in ω from 1 (full indexation) to 0 (no indexation) produces a factual fall in inflation persistence, and I would be back to the standard model with no indexation. The model is indeed successful in reducing persistence. The natural question is then: what is ω? Does a fall from 1 to 0 makes sense? In the backward-looking NK model, a firm i that is unable to reset (log) prices gets to reset its price to

$$p_{it} = p_{it-1} + \omega\pi_{t-1} \tag{G.29}$$

The presence of the term $\omega\pi_{t-1}$ is what gives anchoring. What is the value of ω in the literature? Christiano et al. (2005) assume $\omega = 1$. Smets and Wouters (2007) estimate a value of $\omega = 0.21$ trying to match aggregate anchoring in inflation dynamics. The main problem here is that it is hard to justify a particular micro estimate for ω, since it is unobservable in the micro data. One would need to identify the firms that were not hit by the Calvo fairy in a given period and then regress (G.29). However, the price indexation suggests that all prices are changed in every period, which makes unfeasible to identify the Calvo-restricted firms. Another aspect in which $\omega > 0$ is inconsistent with the micro-data is that it implies that all prices change every period, in contradiction with Bils and Klenow (2004) and Nakamura

and Steinsson (2008). As a result, one cannot claim that ω is the causant of the fall in inflation persistence, since it needs to be identified from the macro aggregate data, which makes unfeasible to separately identify ω and the true inflation persistence.

I therefore conclude that extending the benchmark framework to price indexation does not have the quantitative bite to explain the fall in inflation persistence, although the estimates move in the correct direction.

G.2.3 Trend Inflation

Although it is well known that Central Banks' objective is to have a stable inflation rate around 2%, most New Keynesian models are log-linearized around a zero inflation steady state since the optimal steady state level of inflation is 0%. Ascari and Sbordone (2014) extend the benchmark model to account for trend inflation. The non-linear model is identical to the one presented in the previous section. Differently from the standard environment, they log-linearize the model around a steady with a certain level of trend inflation $\bar{\pi}$, which is constant over time. Price dispersion, a backward-looking variable that has no first-order effects in the benchmark NK model, is now relevant for the trend NK model. Augmenting the model with trend inflation creates intrinsic persistence in the inflation dynamics through relative price dispersion. The model, similar to the one in Ascari and Sbordone (2014), is derived in Online Appendix F.2. The model can now be summarized as a system of six equations, including (3.6), (3.7) and (3.8), with the additional inclusion of the price dispersion dynamics (G.30)

$$s_t = \frac{\epsilon}{1-\alpha}\frac{\delta-\chi}{1-\chi}\pi_t - \frac{\omega\epsilon}{1-\alpha}\frac{\delta-\chi}{1-\chi}\pi_{t-1} + \delta s_{t-1} \tag{G.30}$$

and the Phillips curve, which is now given by the system

$$\begin{aligned}
\pi_t &= \kappa_\pi \pi_{t-1} + \kappa_\psi \psi_t + \kappa_y y_t + \beta_\psi \mathbb{E}_t \psi_{t+1} + \beta_\pi \mathbb{E}_t \pi_{t+1} \\
\psi_t &= (1-\beta\delta)\varphi s_t + \frac{1+\varphi}{1-\alpha}(1-\beta\delta) y_t - \frac{\omega\epsilon}{1-\alpha}\beta\delta\pi_t + \beta\delta\mathbb{E}_t\psi_{t+1} + \frac{\epsilon}{1-\alpha}\beta\delta\mathbb{E}_t\pi_{t+1}
\end{aligned} \tag{G.31}$$

where $\Theta = \frac{1-\alpha}{1-\alpha+\epsilon\alpha}$, $\delta(\overline{\pi}) = \theta\overline{\pi}^{\frac{\epsilon(1-\omega)}{1-\alpha}}$ and $\chi(\overline{\pi}) = \theta\overline{\pi}^{(\epsilon-1)(1-\omega)}$, $\kappa_\pi = \frac{\omega}{1-\omega[\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi]}$, $\kappa_\psi = \frac{\Theta(1-\chi)}{\chi\{1-\omega[\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi]\}}$, $\kappa_y = -\frac{\Theta(1-\sigma)(1-\chi)(1-\beta\chi)}{\chi\{1-\omega[\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi]\}}$, $\beta_\psi = -\frac{\Theta\beta(1-\chi)}{1-\omega[\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi]}$ and $\beta_\pi = -\frac{\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi}{1-\omega[\Theta(\epsilon-1)\beta(1-\chi)-\beta\chi]}$. The parameterization is identical to that of Tables OA.1 and OA.7, extended to trend inflation between 0%-6%, except for the value of $\phi_y = 0$ which is bounded from above by the determinacy conditions. The model can be collapsed to a system

Parameter	Description	Value	Source/Target
$\bar{\pi}$	Trend inflation	$1.02^{1/4} - 1.04^{1/4}$	Ascari and Sbordone (2014)

Table OA.7: Model Parameters

of four second-order stochastic difference equations

$$\boldsymbol{x}_t = \boldsymbol{\Gamma_b} \boldsymbol{x}_{t-1} + \boldsymbol{\Gamma_f} \mathbb{E}_t \boldsymbol{x}_{t+1} + \boldsymbol{\Lambda} v_t$$

where $\boldsymbol{x}_t = [y_t \quad \pi_t \quad \psi_t \quad s_t]'$. The solution of the above system satisfies

$$\boldsymbol{x}_t = A \boldsymbol{x}_{t-1} + \Psi v_t \tag{G.32}$$

where both matrices $A(\phi_\pi, \bar{\pi}, \Phi)$ and $\Psi(\phi_\pi, \bar{\pi}, \Phi)$ depend now on ϕ_π, trend inflation $\bar{\pi}$, and the rest of the model parameters Φ.

In this framework, I define s_t as (log) price dispersion at time t, and ψ_t as the present discounted value of future marginal costs. Notice that I have extended an otherwise standard trend-inflation NK model with price indexation (governed by ω) as in (G.29). Even in the zero-indexation case, there will be anchoring coming from the price dispersion equation, which is the only backward-looking equation in the system. To see this, under zero-indexation, inflation dynamics are given by

$$\begin{aligned} \pi_t &= a_s s_{t-1} + b_\pi v_t \\ &= \left(a_s \frac{\epsilon}{1-\alpha} \frac{\delta - \chi}{1-\chi} + \delta \right) \pi_{t-1} + b_\pi (v_t - \delta v_{t-1}) \end{aligned}$$

In the price-indexation case, inflation dynamics are given by

$$\begin{aligned} \pi_t &= a_\pi \pi_{t-1} + a_s s_{t-1} + b_\pi v_t \\ &= \left(a_\pi + \delta + a_s \frac{\epsilon}{1-\alpha} \frac{\delta - \chi}{1-\chi} \right) \pi_{t-1} - \left(a_\pi \delta + a_s \frac{\omega \epsilon}{1-\alpha} \frac{\delta - \chi}{1-\chi} \right) \pi_{t-2} + b_\pi (v_t - \delta v_{t-1}) \end{aligned}$$

Most importantly, one can see that the parameter that governs anchoring (and persistence) in the system, δ in (G.30), is increasing in the level of trend inflation $\overline{\pi}$. This framework, therefore, has the potential of explaining the fall in inflation persistence *if* trend inflation had fallen. Stock and Watson (2007) and Ascari and Sbordone (2014) provide evidence of a fall of trend inflation from 4% in the 1969-1985 period to 2% afterwards. They estimate trend

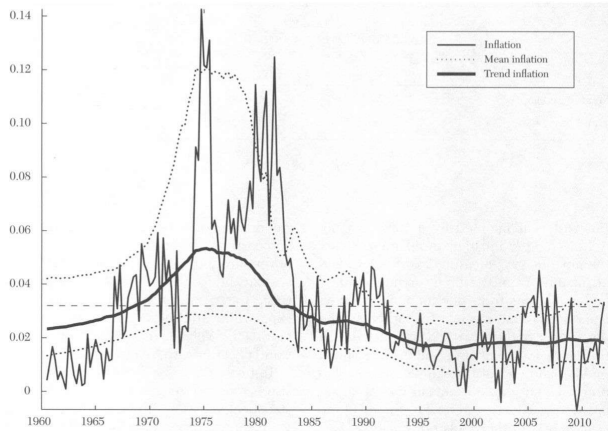


Figure OA.4: Inflation, Trend Inflation and Mean Inflation, Figure 3 in Ascari and Sbordone (2014).

inflation using a Bayesian VAR with time-varying coefficients, which I reproduce here in Figure OA.4. Importantly, they find that their estimated trend inflation is correlated (0.96) with the 10-year inflation expectations reported in the Survey of Professional Forecasters (after 1981).

As I argued before, a fall in the trend inflation $\bar{\pi}$ would decrease $\delta(\bar{\pi})$ and thus reduce aggregate anchoring in the system. I therefore investigate if such fall, together with the already discussed change in ϕ_π, can explain the documented fall in inflation persistence.

I compute the first-order autocorrelation of inflation for values of $(\phi_\pi, \overline{\pi}) \in [1.2, 2] \times [0\%, 6\%]$ in the trend inflation model with price indexation. I plot our results in Figure OA.5. As I previewed above, the decrease in trend inflation documented by Ascari and Sbordone (2014) can explain (part of) the fall in persistence. In particular, a fall in trend inflation from 6% to 2% (holding $\phi_\pi = 1.5$ constant) produces a fall in inflation persistence from 0.887 to 0.851. Similarly, an increase in the aggressiveness towards inflation from 1 to 2 (Clarida et al. 2000), holding $\overline{\pi} = 2\%$ constant, produces a fall in inflation persistence from 0.879 to 0.845. Jointly, they produce a fall from 0.912 to 0.845. Although in the correct direction, the trend inflation model lacks the enough quantitative bite to produce the large fall documented in Table **??**. I therefore conclude that extending the benchmark framework to trend inflation and price indexation does not explain the fall in inflation persistence, although the estimates move in the correct direction.

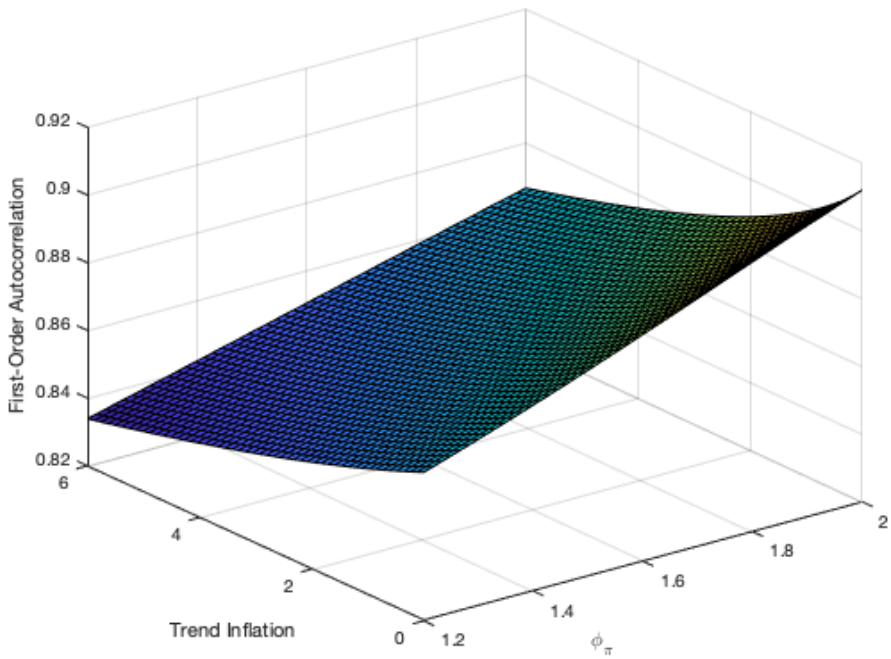


Figure OA.5: First-order autocorrelation for values $(\phi_\pi, \overline{\pi}) \in [1.2, 2] \times [0\%, 6\%]$

H Useful Mathematical Concepts

H.1 Wiener-Hopf Filter

Consider the non-causal prediction of $f_t = A(L)\widehat{\boldsymbol{s}}_{it}$ given the whole stream of signals

$$\begin{aligned}
\mathbb{E}(f_t|x_i^\infty) &= \rho_{yx}(L)\rho_{xx}^{-1}(L)x_{it} \\
&= \rho_{yx}(L)\boldsymbol{B}(L^{-1})^{-1}\boldsymbol{V}^{-1}\boldsymbol{B}(L)^{-1}x_{it} \\
&= \rho_{yx}(L)\boldsymbol{B}(L^{-1})^{-1}\boldsymbol{V}^{-1}\boldsymbol{w}_{it} \\
&= \sum_{k=-\infty}^{\infty} h_k \boldsymbol{w}_{it-k}
\end{aligned}$$

where $\rho_{yx}(z) = A(z)\boldsymbol{M}'(z^{-1})$ and $\rho_{xx}(z) = \boldsymbol{B}(z)\boldsymbol{V}\boldsymbol{B}'(z^{-1})$. Notice that we are using future values of $\boldsymbol{w}_{it}$. However, if the agent only observes events or signals up to time t, the best

prediction is

$$
\begin{aligned}
\mathbb{E}(f_t|x_i^t) &= \left[\sum_{k=-\infty}^{\infty} h_k \boldsymbol{w}_{it-k}\right]_+ \\
&= \sum_{k=0}^{\infty} h_k \boldsymbol{w}_{it-k} \\
&= \left[\rho_{yx}(L)\boldsymbol{B}(L^{-1})^{-1}\right]_+ \boldsymbol{V}^{-1}\boldsymbol{w}_{it} \\
&= \left[\rho_{yx}(L)\boldsymbol{B}(L^{-1})^{-1}\right]_+ \boldsymbol{V}^{-1}\boldsymbol{B}(L)^{-1}x_{it}
\end{aligned}
$$

H.2 Annihilator Operator

The annihilator operator $[\cdot]_+$ eliminates the negative powers of the lag polynomial:

$$
[A(z)]_+ = \left[\sum_{k=-\infty}^{\infty} a_k z^k\right]_+ = \sum_{k=0}^{\infty} a_k z^k
$$

Suppose that we are interested in obtaining $[A(z)]_+$, where $A(z)$ takes this particular form, $A(z) = \frac{\phi(z)}{z-\lambda}$ with $|\lambda| < 1$, and $\phi(z)$ only contains positive powers of z. We can rewrite $A(z)$ as

$$
A(z) = \frac{\phi(z) - \phi(\lambda)}{z-\lambda} + \frac{\phi(\lambda)}{z-\lambda}
$$

Let us first have a look at the second term, We can write

$$
\begin{aligned}
\frac{\phi(\lambda)}{z-\lambda} &= -\frac{\phi(\lambda)}{\lambda}\frac{1}{1-\lambda^{-1}z} \\
&= -\frac{\phi(\lambda)}{\lambda}(1 + \lambda^{-1}z + \lambda^{-2}z^2 + ...)
\end{aligned}
$$

which is not converging. Alternatively, we can write it as a converging series as

$$
\begin{aligned}
\frac{\phi(\lambda)}{z-\lambda} &= \phi(\lambda)z^{-1}\frac{1}{1-\lambda z^{-1}} \\
&= \phi(\lambda)z^{-1}(1 + \lambda z^{-1} + \lambda^2 z^{-2} + ...)
\end{aligned}
$$

Notice that all the power terms are on the negative side of z. As a result,

$$\left[\frac{\phi(\lambda)}{z-\lambda}\right]_+ = 0$$

Let us now move to the first term. We can write

$$\begin{aligned}\phi(z) - \phi(\lambda) &= \sum_{k=0}^{\infty} \phi_k (z^k - \lambda^k) \\ &= \phi_0 \prod_{k=1}^{\infty} (z - \xi_k)\end{aligned}$$

where $\{\xi^k\}$ are the roots of this difference polynomial. Since we know that λ is a root of the LHS, we can set $\xi^k = \lambda$ and write

$$\phi(z) - \phi(\lambda) = \phi_0 (z-\lambda) \prod_{k=2}^{\infty} (z - \xi_k) \implies \frac{\phi(z) - \phi(\lambda)}{z-\lambda} = \prod_{k=2}^{\infty} (z - \xi_k)$$

which only contains positive powers of z. Hence, we have that

$$\left[\frac{\phi(z)}{z-\lambda}\right]_+ = \frac{\phi(z) - \phi(\lambda)}{z-\lambda}$$

Consider now instead the case $A(z) = \frac{\phi(z)}{(z-\lambda)(z-\beta)}$. Making use of partial fractions, we can write

$$\begin{aligned}\frac{\phi(z)}{(z-\lambda)(z-\beta)} &= \frac{1}{\lambda-\beta}\left[\frac{\phi(z)}{z-\lambda} - \frac{\phi(z)}{z-\beta}\right] \\ &= \frac{1}{\lambda-\beta}\left[\frac{\phi(z)-\phi(\lambda)}{z-\lambda} - \frac{\phi(z)-\phi(\beta)}{z-\beta} + \frac{\phi(\lambda)}{z-\lambda} - \frac{\phi(\beta)}{z-\beta}\right]\end{aligned}$$

Following the same steps as in the previous case, we can solve

$$\left[\frac{\phi(z)}{(z-\lambda)(z-\beta)}\right]_+ = \frac{\phi(z)-\phi(\lambda)}{(\lambda-\beta)(z-\lambda)} - \frac{\phi(z)-\phi(\beta)}{(\lambda-\beta)(z-\beta)}$$