Inflation Inattention on the Production Network: Firm-Level Evidence and Macro Implications

Rafael Gonçalves Ina Hajdini

February 2026

Click here for the latest version

Abstract

We study the determinants and macroeconomic consequences of heterogeneity in firms' inflation expectations, focusing on the role of industry differences. Relying on a micro-level dataset of US firms' inflation expectations, we document that firms in industries with larger Domar weights and more flexible prices have more accurate inflation forecasts. To explain these patterns, we develop a rational inattention model of price-setting firms within a production network. In the model, inattention compounds downstream through input-output linkages, and firms' forecast accuracy depends on their attention to marginal costs and the comovement between marginal costs and aggregate inflation. When calibrated to US input-output data, the model replicates the positive cross-industry relationship between forecast accuracy, Domar weights, and price flexibility. Quantitatively, we show that a selection effect of rational inattention steepens the Phillips curve relative to the exogenous-attention benchmark, partially offsetting the standard flattening effect of input-output linkages.

Keywords: inflation expectations, rational inattention, production networks, inflation dynamics, monetary policy.

JEL codes: E31, E32, E71.

Gonçalves: Princeton University, rgoncalves@princeton.edu. Hajdini: Federal Reserve Bank of Cleveland, Ina.Hajdini@clev.frb.org. We thank Karthik Sastry, Gianluca Violante, and participants at PUC-Rio, the 2025 Paris Conference on the Macroeconomics of Expectations, the Spring 2025 Midwest Macroeconomics Meeting, and CEBRA Annual Meeting for helpful comments. The views expressed here are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

1. Introduction

Modern macroeconomic theory attributes a crucial role to firms' inflation expectations in shaping the dynamics of both nominal and real variables. However, limited data on firms' expectations have hindered our understanding of the joint distribution of firms' price decisions, their beliefs, and macroeconomic outcomes. In this context, the full-information rational expectations (FIRE) framework has served as the paradigm for modeling expectations and their influence on macroeconomic dynamics for the past decades.

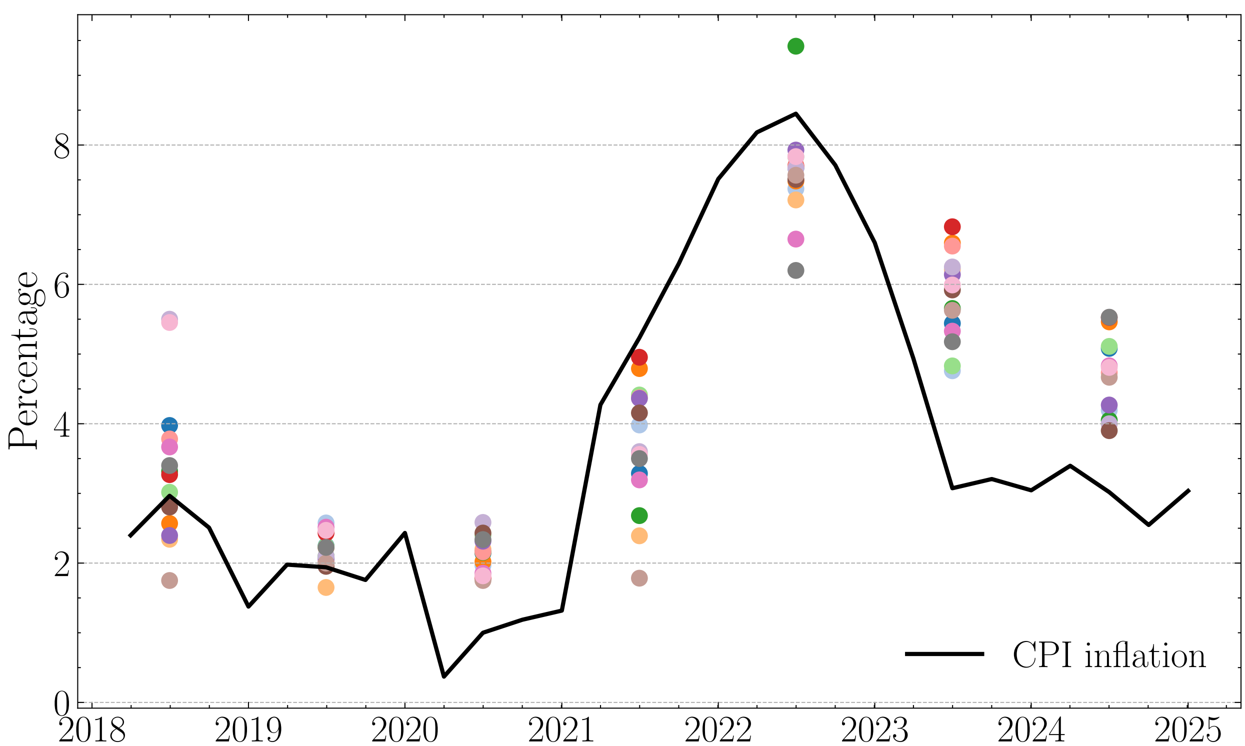
A growing empirical literature has documented this heterogeneity and inaccuracy in firms' inflation expectations (Coibion, Gorodnichenko, and Kumar 2018; Candia, Coibion, and Gorodnichenko 2024, among others). Notably, firms' "inattention" to inflation appears related to industry-specific characteristics, such as experienced input price changes (Andrade et al. 2022) and the number of competitors (Afrouzi 2024). However, the role of other relevant industry attributes—such as an industry's importance in the production network and the industry-level degree of price rigidity—in shaping firms' expectations remains underexplored.

This paper investigates the causes and macroeconomic consequences of heterogeneity in firms' inflation expectations, with a focus on the role of industry differences. On the empirical front, we combine a novel and comprehensive dataset on US firms' inflation expectations with various data sources on industry characteristics. We document that firms' forecast accuracy of inflation is systematically associated with industry attributes: firms in industries with larger Domar weights (sales-to-GDP ratios) and more flexible prices have significantly more accurate inflation forecasts. We also find supportive evidence that firms form inflation expectations based on the realized industry-specific marginal costs, consistent with the view that firms use their own cost conditions as a signal about aggregate inflation.

Motivated by these empirical patterns, we develop a rational inattention model of price-setting firms operating within a production network. We first analyze a static special case to derive analytical results. On the determinants of heterogeneous attention, two results emerge: inattention compounds downstream through the network; and firms' inflation forecast accuracy depends on both their attention to industry-specific marginal costs—the relevant variable for price setting—and the comovement between marginal costs and aggregate inflation. On the macroeconomic implications, heterogeneous attention alters the aggregate effects of shocks, and the relationship between the slope of the Phillips curve and the strength of input-output linkages becomes ambiguous. We then embed the model in a dynamic setting and calibrate it to the US economy to quantify these forces and their implications for inflation dynamics and monetary policy.

Evidence. Our empirical analysis draws on the Survey of Firms' Inflation Expectations (SoFIE), a nationally representative quarterly survey of US firms, which we combine with industry-level data on production structure and pricing behavior. Figure 1 plots industry-average perceived inflation rates against the actual CPI inflation rate over time. The dispersion is striking: firms in different industries hold systematically different beliefs about recent inflation, with some industries tracking the actual rate closely and others deviating substantially.

FIGURE 1. Perceptions about inflation across industries



Note: The black line is the real-time inflation rate from the Federal Reserve Bank of Philadelphia Real-Time Data Research Center. The colored dots are average perceived inflation rates within the following industries: Basic Metals, Electrical, Financial Intermediation, Food and Drink, Hotels and Restaurants, Mechanical Engineering, Other Manufacturing, Other Services, Post and Telecommunication, Renting and Business Activities, Timber and Paper, Transport Manufacturing, Transport and Storage Services. Sample from 2Q2018 to 1Q2024.

To investigate what drives this cross-industry dispersion, we combine the SoFIE with data on input-output linkages and industry-level price rigidity. We measure firms' attention to inflation using the absolute value of forecast errors and regress these on industry characteristics, including the industry's Domar weight, upstreamness, price flexibility, and the volatility of productivity shocks. We find that firms in industries with larger Domar weights and more flexible prices have significantly more accurate inflation forecasts, while upstreamness and TFP volatility do not have significant effects once we control for these variables.

Model. To rationalize our empirical findings and study their macroeconomic implications, we develop a dynamic rational inattention model of price-setting firms operating within a production network. Firms produce differentiated goods using labor and intermediate inputs sourced from other industries. The economy is subject to aggregate nominal demand shocks and industry-specific productivity shocks, and firms set prices under incomplete information about these shocks while optimally choosing how to allocate their limited attention. Crucially, the information friction is the source of both heterogeneous expectations and nominal rigidities, so attention and price flexibility are jointly determined.

To build intuition, we first specialize the model to the case where agents are myopic and shocks are i.i.d., which yields a tractable static environment with closed-form results. This simplified setting delivers the key analytical insights described below. We then return to the general dynamic model for quantitative analysis.

Equilibrium attention. To set prices optimally, firms need to form beliefs about their industry-specific marginal cost while considering the costs of acquiring information about the state of the economy. Equilibrium firms' attention is determined by (i) the cost of obtaining information, (ii) the profit losses due to a pricing mistake of a given size, and (iii) the volatility of marginal costs. Importantly, the second and third components are industry-specific, implying that firms in different sectors have varying incentives to acquire information.

Downstream compounding of inattention (and price rigidity). At the heart of the model is a fixed point between the responsiveness of marginal costs to shocks and attention levels that determine nominal rigidities. Intuitively, low attention in a certain industry results in dampened responses of that industry's prices to aggregate and sectoral shocks. This leads to a reduction in the volatility of marginal costs in downstream sectors, thus reducing the benefits of acquiring information and encouraging further inattention within those sectors. This intuition leads to our first key theoretical result: inattention compounds downstream due to strategic complementarities in pricing and information acquisition decisions.

The determinants of inflation expectations. Our framework yields a decomposition of firms' inflation expectations into three components: (i) firms' attention to their industry-specific marginal cost, (ii) the comovement between marginal costs and aggregate inflation, and (iii) changes in marginal costs. This decomposition arises because firms primarily acquire information about their marginal cost, the relevant variable for price setting, and use this imperfect knowledge to form beliefs about aggregate inflation. A direct corollary connects forecast accuracy to industry characteristics: forecast accuracy about aggregate inflation is increasing in both the firm's attention level and the R-squared of a regression of aggregate inflation on industry-level marginal cost changes. Thus, forecast accuracy depends not only on how attentive firms are, but also on how informative their industry's cost conditions are about aggregate inflation.

Heterogeneous attention and the aggregate effects of shocks. We now turn to the aggregate implications of heterogeneous attention. The aggregate price response to any shock decomposes into two channels: a representative-agent channel, driven by the average level of attention, and a heterogeneous-attention channel, captured by the covariance between industries' attention levels and their marginal cost exposure to the shock. When more attentive industries are also those whose marginal costs respond more strongly to the shock, heterogeneous attention amplifies the aggregate price response; the reverse holds when less exposed industries are the most attentive. This decomposition highlights that not only the average degree of inattention matters for macroeconomic outcomes, but also the cross-sectional pattern of which industries are attentive.

The slope of the Phillips curve. The production network also shapes the slope of the Phillips curve, but the direction depends on whether attention is endogenous. Under exogenous nominal frictions, stronger input-output linkages flatten the Phillips curve by compounding rigidities across industries (Basu 1995; Rubbo 2023). Under endogenous attention, a countervailing force arises: changes in input-output linkages influence the volatility of marginal costs, which affects optimal attention levels and the slope of the aggregate Phillips curve. In our calibrated economy, the second force partially offsets the first: the selection effect—whereby high-exposure industries are also the most attentive—yields a steeper Phillips curve than the homogeneous, exogenous attention benchmark.

Quantitative results. To quantify these analytical insights, we calibrate the dynamic model to the US economy using input-output tables, industry-level productivity data, and the SoFIE. We first validate the model against the empirical patterns documented in Section 2. The calibrated model replicates the cross-industry relationship between forecast accuracy and industry characteristics: industries with larger Domar weights and more flexible prices have smaller forecast errors, as in the data. We further test the model's prediction that firms form inflation beliefs through their observation of own marginal costs, and find strong support: firms' inflation expectations are significantly associated with industry-level marginal costs.

We then study the propagation of monetary shocks through the production network. Production networks attenuate the inflation response to monetary policy by roughly 15% while amplifying and prolonging the real output response. Importantly, which industries are attentive matters for aggregate dynamics. In our calibrated economy, industries with large Domar weights have higher equilibrium attention. In the data, these industries also tend to be labor-intensive, which increases their marginal cost exposure to monetary shocks. This positive correlation between attention and exposure amplifies the aggregate inflation response relative to a counterfactual with homogeneous attention.

Related literature. Our model builds on the literature on rational inattention in price-setting problems (Sims 2003; Mackowiak and Wiederholt 2009, among many others), which has mainly abstracted from input-output linkages. Afrouzi (2024) studies how oligopolistic competition shapes firms' strategic inattention to aggregates, showing that firms with fewer competitors are less attentive; Fang et al. (2024) use EDGAR browsing data to measure firm-to-firm attention allocation in production networks, documenting that firms pay more attention to closer and more volatile network neighbors. Our paper complements both by studying attention to aggregate inflation within a production network, providing a theoretical decomposition of CPI expectations and empirical evidence using inflation forecast errors.

Two other recent contributions share a key mechanism with our paper: firms form beliefs about aggregate conditions based on the realization of their own cost conditions. Minton and Monnery (2025) use US survey data on firm-level cost expectations, documenting over-reaction to idiosyncratic cost movements and under-reaction to aggregate shocks. Gagliardone and Tielens (2025) use Belgian micro-data linking firm-level prices, costs, and survey expectations to show that firms respond strongly to realized cost shocks but exhibit muted responses to expected future cost growth. In our framework, this learning mechanism arises endogenously from rational inattention, and the production network structure determines which industries have stronger incentives to attend to their costs.

A growing body of literature examines inflation and monetary policy within production networks (Pastén, Schoenle, and Weber 2024; Rubbo 2023, e.g.,). Our model particularly builds upon La'O and Tahbaz-Salehi (2022)'s approach, which introduces nominal rigidities to a multi-sector economy through incomplete information in a static setting. However, when firms' attention is endogenous to the production structure, conventional results—such as flatter Phillips curves in economies with stronger intermediate input shares—may no longer hold.

On the empirical side, a growing literature documents departures of firms' inflation expectations from full-information rational expectations (FIRE) predictions and explores their determinants (Coibion, Gorodnichenko, and Kumar 2018; Andrade et al. 2022; Candia, Coibion, and Gorodnichenko 2024; Weber et al. 2025). Notably, Candia, Coibion, and Gorodnichenko (2024), also using the SoFIE, finds systematic differences in firms' forecasts across industries. In related work, Hajdini et al. (2025) use a randomized control trial of firm pairs in New Zealand to show that macroeconomic expectations propagate along supply chain networks through firm-to-firm communication. To the best of our knowledge, our paper is the first to link micro-level data on firms' inflation expectations to sources of industry attributes, guided by new theory.

Finally, this paper connects to a large literature that studies the implications of incomplete information to inflation dynamics and monetary non-neutrality (Lucas 1972; Woodford 2002; Mankiw and Reis 2002; Nimark 2008; Angeletos and Huo 2021, among others). As in those papers, the information friction is the source of nominal rigidities and monetary non-neutrality. However, in our model, not only does the average degree of inattention matter for macroeconomic outcomes, but its cross-sectional heterogeneity does as well.

Outline. The paper proceeds as follows. Section 2 presents our data and our new empirical results relating US firms' inflation expectations to industry characteristics. In Section 3, we introduce the general dynamic rational inattention, multi-sector model of price setting. Section 4 characterizes the equilibrium and derives analytical results for a simplified version of the model. In Section 5, we present the model's theoretical insights on the determinants and macroeconomic implications of heterogeneity in firms' inflation expectations. Section 6 presents quantitative results from the calibrated general model. We evaluate the model's fit to the evidence on forecast accuracy and industry characteristics, and then we study the implications for monetary policy. Section 7 concludes.

2. Empirical Results

In this section, we present a number of stylized facts relating the accuracy of firms' inflation expectations to their industry characteristics.

2.1. Data

The Survey of Firms' Inflation Expectations (SoFIE). At the center of our empirical analysis is the firm-level data on inflation expectations sourced from the Survey of Firms' Inflation Expectations (SoFIE). The SoFIE is a nationally representative, quarterly survey of inflation expectations conducted by the Federal Reserve Bank of Cleveland. It elicits inflation expectations from CEOs and other top executives and has been conducted since the second quarter of 2018. The survey relies on the recurring participation of a panel of firms, with between 300 and 700 firms participating each quarter. Importantly, the dataset provides information about each firm, including the sector in which the firm operates (services or manufacturing), its industry, and its size (measured by a categorical classification based on the number of employees).[1]

We make use of two questions about beliefs about aggregate inflation in the survey. The first question — asked every quarter — asks firms to provide their expectations for the Consumer Price Index (CPI) inflation rate over the upcoming twelve-month period:

> *What do you think will be the inflation rate (for the Consumer Price Index) over the next 12 months? Please provide an answer in an annual percentage rate.*

The second question we rely on — elicited every third quarter of the year — asks about their perceptions of what the inflation rate has been over the last twelve months:

[1] Firms are selected randomly from the manufacturing and services sectors to accurately represent the underlying structure of each sector's contribution to aggregate gross value added. Within the manufacturing sector, companies are classified into food and drink, textiles and clothing, electrical, chemicals and plastics, transport, timber and paper, basic metals, mechanical engineering, and other manufacturing. Within the services sector, the companies are classified into hotels and restaurants, transport and storage, post and telecommunication, financial intermediation, renting and business activities, and other services. The firm size classification is categorical: firms are divided into small (1-19 employees), medium (20-249 employees), and large (250+ employees).

What do you think has been the annual inflation rate (for the Consumer Price Index) over the last twelve months? Please provide an answer in annual percentage rate.

See Appendix C.1 for additional details on the survey methodology and question rotation.

Industry-level data sources. We use three data sources with industry-level information to empirically study how industry characteristics explain differences in firms' inflation expectations. First, we use the 2022 input-output tables constructed by the Bureau of Economic Analysis (BEA) to compute different measures of industry importance within the production network or the overall economy. Our second data source is from Pastén, Schoenle, and Weber (2024), who compute measures of the probability of price adjustments at the industry level using the microdata that underlies the Bureau of Labor Statistics (BLS) producer price index (PPI). They construct this measure by taking the ratio of the number of price changes to the number of sample months. Third, we use the BEA/BLS Integrated Industry-Level Production Account (ILPA) data, which provides estimates of industry-specific productivity from 1987 to 2021. Table A1 in the appendix provides additional details on data sources and variable definitions.

2.2. Measurement

Forecast errors. We calculate the accuracy of firms' inflation forecasts as the absolute difference between one-year-ahead forecasts and the actual (12-month-ahead) realizations of US consumer price index inflation. We start by constructing a series of firms' inflation forecast errors as the difference between the actual CPI inflation rate and the firm's forecast:

$$\mathrm{FE}_{fit}[\pi_{t+h}] \equiv \pi_{t+h} - \mathbb{E}_{fit}[\pi_{t+h}], \quad \text{for } h \in \{0, 4\}. \tag{1}$$

Figure A1 in the appendix shows that the time series of industry inflation forecasts exhibit similar dynamics in levels, closely aligning with inflation fluctuations. However, some industries exhibit significantly more volatile expectations compared to others.

Industry importance. Measures of industries' importance in the economy are constructed at the two-digit ISIC industry classification level for consistency with the SoFIE dataset. Our calculations follow the construction of direct input and labor requirements in Rubbo (2023).

Specifically, we compute industries' total sales as a fraction of GDP, also known as sales-based Domar weights, and a measure of industry "upstreamness" based on Antràs et al. (2012), which captures the importance of a particular industry as a *direct and indirect* supplier to all industries in the economy. Formally, let A be the input-output matrix and $L \equiv (I - A)^{-1}$ be the Leontief inverse matrix. The entry (i, j) of the Leontief inverse matrix L quantifies the direct and indirect dependencies of sector i on sector j. The upstreamness measure is given by

$$u_i \equiv \sum_{k=1}^{n} \ell_{ki} - \mathbb{I}\{i = k\}, \tag{2}$$

i.e., it is the i-th column sum of the Leontief inverse (Carvalho and Tahbaz-Salehi 2019).

TFP volatility. We compute the volatility of industry-level productivities after linearly detrending the productivity series from ILPA for each industry.

Industry-specific price stickiness. We aggregate Pastén, Schoenle, and Weber (2024)'s estimates of industry-level probabilities of price readjustment to match the two-digit industry classification in the SoFIE data by calculating a simple average based on the more granular industry classification they have.

2.3. Results

We estimate how industry characteristics correlate with the accuracy of firms' inflation forecasts by running the following regression:

$$|\text{FE}_{fit}[\pi_{t+h}]| = \chi_t + \boldsymbol{\gamma}'\boldsymbol{X}_i + \boldsymbol{\delta}'\boldsymbol{D}_{fit} + \text{error}_{fit}, \quad \text{for } h \in \{0, 4\} \tag{3}$$

where $|\text{FE}_{fit}[\pi_{t+h}]|$ is the absolute inflation forecast error about h=quarters-ahead inflation of firm f in industry i in quarter t; χ_t denotes time fixed effects. $\boldsymbol{X}_i$ is a vector of time-invariant industry characteristics: the industry Domar weight (sales-to-GDP ratio), industry upstreamness, price flexibility (the industry-specific probability of price readjustment), and the volatility of industry-specific productivity shocks. $\boldsymbol{\gamma}$ are the coefficients of interest. $\boldsymbol{D}_{fit}$ is a vector of firm size dummies (medium, large; small is the reference category). Table 1 summarizes the results of this regression.

The coefficient on the Domar weight in regression (3) is negative and statistically significant at the 1% level, indicating that firms in industries with larger Domar weights have more accurate inflation nowcasts. The estimated coefficient of −0.831 (column 3) implies

TABLE 1. Forecast Errors and Industry Characteristics

	$\lvert FE_{fit}[\pi_t]\rvert$			$\lvert FE_{fit}[\pi_{t+4}]\rvert$		
	(1)	(2)	(3)	(4)	(5)	(6)
Domar weight	-0.730***	-0.882***	-0.831***	-0.014	-0.260**	-0.218*
	(0.248)	(0.234)	(0.234)	(0.143)	(0.119)	(0.119)
Upstreamness	0.026	0.029	0.025	-0.012	-0.003	-0.009
	(0.043)	(0.040)	(0.040)	(0.023)	(0.019)	(0.019)
TFP volatility	0.001	0.007	0.003	-0.013	-0.006	-0.008
	(0.033)	(0.031)	(0.031)	(0.021)	(0.018)	(0.018)
Price flexibility	-3.487***	-3.816***	-3.635***	-0.505	-1.208**	-1.002**
	(1.020)	(0.963)	(0.961)	(0.577)	(0.473)	(0.475)
Observations	2,178	2,178	2,178	7,884	7,884	7,884
R^2	0.021	0.136	0.139	0.002	0.305	0.309
Time FE	No	Yes	Yes	No	Yes	Yes
Size controls	Yes	No	Yes	Yes	No	Yes

Note: This table reports regressions of absolute forecast errors of inflation on industry characteristics. Size controls correspond to firm size dummies (Medium, Large; reference: Small). Robust standard errors in parentheses. ***, **, * indicate statistical significance at 1, 5, and 10 percent, respectively.

that a one-percentage-point increase in an industry's sales-to-GDP ratio is associated with a 0.83 percentage point reduction in absolute nowcast errors. The coefficient on price flexibility is also negative and highly statistically significant. Since price flexibility is measured as the probability of price readjustment, the coefficient of −3.635 implies that a 10 percentage point increase in an industry's frequency of price adjustment is associated with a 0.36 percentage point reduction in absolute nowcast errors. Across the full range of price flexibility in our sample, the implied difference is approximately 2.6 percentage points.

In contrast, we find that upstreamness and TFP volatility do not have statistically significant effects on forecast accuracy once we control for Domar weights and price flexibility. In Appendix C.4, we show that these results are robust to including alternative measures of industry importance, such as cost-based Domar weights.

We regress firms' inflation expectations on our measure of industry-specific marginal

TABLE 2. Inflation expectations and marginal costs

	(1)	(2)	(3)	(4)	(5)
	Dependent variable:				
	1-year ahead forecast			Nowcast	
mc_{it}	0.021***	0.020***	0.016**	0.019***	0.018***
	(0.0053)	(0.0072)	(0.008)	(0.0067)	(0.0071)
Time FE		✓	✓		✓
Firm FE			✓		
Observations	7,884	7,884	7,884	2,178	2,178
R-squared	0.321	0.395	0.486	0.310	0.385

Note: The table shows the results of regression (4). For columns 1-4, the outcome is the firms' 12-month-ahead inflation expectations. For columns 5-6, the outcome is the firm's inflation nowcast. Standard errors are clustered at the industry level. ***, **, * indicate statistical significance at 1, 5, and 10 percent, respectively.

costs:

$$\mathbb{E}_{fit}[\pi_t] = \chi_f + \chi_t + \beta \mathrm{mc}_{it} + \mathrm{error}_{fit}. \tag{4}$$

We guard against potential misspecifications in our model by including firm and time fixed effects. Firm-level fixed effects control for unobserved heterogeneity in firms' forecasting abilities, while time fixed effects capture aggregate shocks that may simultaneously affect both marginal costs and inflation expectations.

The regression results in Table 2 support the key prediction of the model: firms' inflation expectations are significantly correlated with their industry-specific marginal costs. The estimated coefficient on mc_{it} ranges from 0.016 to 0.021 for 12-month-ahead forecasts and from 0.018 to 0.019 for nowcasts, and is statistically significant at the 1% level in nearly all specifications. This evidence is consistent with the decomposition in Proposition 2, which predicts that firms learn about aggregate inflation indirectly through their observation of own marginal costs.

Having documented these cross-industry patterns, we now develop a structural model that rationalizes them and allows us to study their macroeconomic implications. The results in Table 1 highlight that firms' inflation expectations are not consistent with the full information rational expectations hypothesis, and that firms' performance in terms of forecasting inflation is highly dependent on the position of the firm in the production network. To rationalize the latter, we next build a multi-sector model of rationally inattentive price-setting firms operating within a production network.

3. Model

Motivated by the evidence presented in the previous section, we build a multi-sector model of rationally inattentive price-setting firms.

3.1. Environment

Time is discrete and indexed by $t \in \{0, 1, \ldots\}$.

Intermediate firms. There are n industries in the economy, indexed by $i \in I \equiv \{1, 2, \ldots, n\}$. Each industry comprises a unit mass of monopolistically competitive firms indexed by $f \in [0, 1]$. The output of each industry can be consumed by households or used as an intermediate input for production by firms in all industries, including the same industry.

The monopolistically competitive firms within each industry set prices and use a common constant-returns-to-scale, Cobb-Douglas technology F_i to transform labor and intermediate inputs into differentiated goods:

$$\begin{aligned} Y_{fit} &= Z_{it} F_i(L_{fit}, X_{fit,1}, \ldots, X_{fit,n}) \\ &= Z_{it} \zeta_i L_{fit}^{\alpha_i} \prod_{k=1}^{n} X_{fit,k}^{a_{ik}}, \quad \alpha_i + \sum_{k=1}^{n} a_{ik} = 1, \end{aligned} \tag{5}$$

where Y_{fit} is the firm's output, L_{fit} is its labor input, and $X_{fit,k}$ is the amount of sectoral commodity $k \in I$ purchased by the firm. Additionally, Z_i is an industry-level productivity shock, and $\zeta_i \equiv \alpha_i^{-\alpha_i} \prod_k a_{ik}^{-a_{ik}}$ is a normalization constant that simplifies the calculations. We assume that industry-level productivities are log-normally distributed according to

$$\log \boldsymbol{Z}_t \sim \mathcal{N}(0, \Sigma), \tag{6}$$

where $\boldsymbol{Z}_t \equiv (Z_{it})_{i \in I}$ and Σ is the covariance matrix. The economy's input-output linkages are summarized by the matrix $A \equiv [a_{ij}]_{i,j \in I}$ and industry labor shares by the vector $\boldsymbol{\alpha} \equiv (\alpha_i)_{i \in I}$.

Firms' nominal profits are given by

$$\Pi_{fit} = (1 - \tau_i) P_{fit} Y_{fit} - W_t L_{fit} - \sum_{k=1}^{n} P_{kt} X_{fit,k}, \tag{7}$$

where P_{fit} is the nominal price charged by the firm, P_{kt} is the nominal price of industry k's sectoral output, and τ_i is an industry-specific revenue tax (or subsidy) levied by the government. Differentiated products produced by the unit mass of firms in each industry i are aggregated into a sectoral good using a constant-elasticity-of-substitution (CES) production technology with elasticity of substitution θ_i:

$$Y_{it} = \left(\int_0^1 Y_{fit}^{(\theta_i - 1)/\theta_i} df \right)^{\theta_i/(\theta_i - 1)}. \tag{8}$$

It follows that firm (i, f)'s demand schedule is given by

$$Y_{fit}^d = \left(\frac{P_{fit}}{P_{it}} \right)^{-\theta_i} Y_{it}. \tag{9}$$

Households. The representative household has Golosov and Lucas (2007)'s preferences defined over a consumption aggregate C and total labor supply L:

$$\max \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\ln C_t - L_t \right) \right], \tag{10}$$

β is the discount factor. The final consumption basket is given by $C_t = \mathcal{C}(C_{1t}, \ldots, C_{nt})$, where C_{it} is the household's consumption of the good produced by industry i and $\mathcal{C}$ is a homogeneous function of degree one. We assume the consumption aggregator $\mathcal{C}$ is a Cobb-Douglas function expressed as follows

$$\mathcal{C}(C_{1t}, \ldots, C_{nt}) = \prod_{i=1}^{n} C_{it}^{\gamma_i}, \quad \sum_{i=1}^{n} \gamma_i = 1. \tag{11}$$

The representative household's budget constraint is given by

$$P_t C_t + B_t \leq W_t L_t + (1 + i_{t-1}) B_{t-1} + \sum_{i=1}^{n} \int_0^1 \Pi_{fit} df + T_t, \tag{12}$$

where $P_t = \mathcal{P}(P_{1t}, \ldots, P_{nt})$ is the nominal price index implied by the consumption aggregator $\mathcal{C}$, B_t are nominal bond holdings, i_t is the nominal interest rate, and T_t are lump-sum government transfers to households.

Wage stickiness. We model wage stickiness a la Blanchard and Galí (2007):

$$W_t = (W_t^*)^{\delta_w} (W_{t-1})^{1-\delta_w} \tag{13}$$

where $\delta_w \in [0, 1]$ denotes the degree of wage flexibility and W_t^* is the nominal wage in the flexible-price equilibrium.

Government. The government's fiscal instrument is a collection of industry-specific taxes (or subsidies) on firms' revenues, with the resulting tax collection then rebated to the representative household as a lump-sum transfer. The government's budget constraint is given by

$$T_t + B_t = (1 + i_{t-1})B_t + \sum_{i=1}^{n} \tau_i \int_0^1 P_{fit} Y_{fit} df. \tag{14}$$

We assume that the central bank controls nominal demand $M_t \equiv P_t C_t$ according to:[2]

$$\Delta \log M_t = \rho_m \Delta \log M_{t-1} + \varepsilon_{mt}, \quad \varepsilon_{mt} \sim \mathcal{N}(0, \sigma_m^2). \tag{15}$$

Information structure. Firms set prices under *rational inattention* about nominal demand M_t and sectoral productivities $\boldsymbol{Z}_t$.

Given their correct prior given by (6) and (15), they choose which signals to observe from a rich set of available signals, $\mathbb{S}$, subject to an information processing constraint. Firms acquire an attention capacity $\kappa_{fit} \geq 0$ subject to a linear cost $\omega \times \kappa_{fit}$. Here, $\omega > 0$ is the information cost parameter.

After firms make their information choices, all shocks and signals are drawn, and each firm observes the realization of its signals. Firms then choose their prices conditional on their information sets, after which demand for each variety is realized. Firms then hire enough labor and intermediate inputs to satisfy demand according to the production function (5). While nominal prices are set under incomplete information, we assume that firms and the representative household make their quantity decisions after observing the prices and the shocks' realizations.

Formally, a strategy for each firm is to choose an information processing capacity κ_{fit}, a set of signals to observe $S_{fit} \subset \mathbb{S}$, and a pricing strategy that maps its information set to a pricing decision, $P_{fit} : S_{fit} \to \mathbb{R}_+$. Given a strategy for all the other firms in the economy,

[2]This is a standard approach to model monetary policy. See Golosov and Lucas (2007), Afrouzi (2024).

firm (i, f) maximizes its profits given its prior:

$$\begin{gathered}\max_{S_{fit}\subset\mathbb{S}, P_{fit}(S_{fit}), \kappa_{fit}} \mathbb{E}\left[(1-\tau_i)P_{fit}Y_{fit}^d - WL_{fit} - \sum_{k=1}^{n} P_k X_{fit,k} - \omega\kappa_{fit}\right] \\ \text{s.t.} \\ Y_{fit}^d = \left(\frac{P_{fit}}{P_{it}}\right)^{-\theta_i} Y_i \quad \text{(demand schedule)} \\ \mathcal{I}(S_{fit}; (M, (P_k)_{k\in I}, \boldsymbol{Z})) \le \kappa_{fit} \quad \text{(information processing constraint)}\end{gathered} \tag{16}$$

where $\mathcal{I}(\cdot;\cdot)$ is Shannon's mutual information function and measures the amount of information that signals S_{fit} contain about the variables that are relevant to firms' pricing decision, namely, nominal demand M_t, industry prices $\boldsymbol{P}_t$, and sectoral productivities $\boldsymbol{Z}_t$.

Equilibrium. An equilibrium for this economy is an allocation for the household,

$$\Omega_H \equiv \{(C_{fit})_{i\in I, f\in[0,1]}, L_t, B_t : \quad t = 0, 1, \ldots\}, \tag{17}$$

a strategy profile for firms given an initial set of signals

$$\Omega_F \equiv \{(\kappa_{fit}, S_{fit} \subseteq \mathbb{S}, P_{fit}, L_{fit}, (X_{fit,k})_{k\in I}, Y_{fit})_{i\in I, f\in[0,1]} : \quad t = 0, 1, \ldots\} \tag{18}$$

and a set of prices

$$\{(P_{fit})_{i\in I, f\in[0,1]}, W_t, i_t : \quad t = 0, 1, \ldots\},$$

such that: (a) given prices and Ω_F, Ω_H solves the household's utility maximization problem; (b) given prices and Ω_H, no firm has an incentive to deviate from Ω_F; (c) $\{M_t \equiv P_t C_t\}$ satisfies the monetary policy rule (15) ; (d) the government budget (14) is balanced; (e) labor and goods markets clear.

4. Analytical Results

To build intuition for the key insights of our framework, we analyze a simplified version of the model in which we set the discount factor $\beta = 0$ and assume $M_t \sim \mathcal{N}(0, \sigma_m^2)$, so that agents are myopic and shocks are i.i.d. We first characterize firms' optimal pricing and information acquisition decisions, and then derive analytical results on the determinants and macroeconomic implications of heterogeneous attention.

4.1. Firms' price-setting and information acquisition decisions

Firms face a decision-making problem comprising four choices: (i) determining attention capacity κ_{if}, (ii) selecting signals to observe S_{if}, (iii) setting the price for their differentiated product p_{if}, and (iv) optimizing production inputs $(\boldsymbol{X}_{fit}, L_{fit})$. We will solve the firms' problem backward, focusing on their information acquisition and pricing decisions.

Inputs choice. When choosing inputs, firms operate under complete information, making decisions once all signals and prices are realized. Their objective at this stage is to minimize costs while meeting the demand for their products Y^d_{fit}, taking wages and input prices as given:

$$\min W_t L_{fit} + \sum_{k=1}^{n} P_{kt} X_{fit,k} \quad s.t.: \quad F_i(L_{fit}, X_{fit,k}) = Y^d_{fit}. \tag{19}$$

Firms' demand for labor and intermediate inputs is given by

$$L_{fit} = Z_{it}^{-1} \alpha_i \left(\prod_{j=1}^{n} P_{jt}^{a_{ij}} \right) W_t^{\alpha_i - 1} Y_{fit} \tag{20}$$

$$X_{fit,k} = Z_{it}^{-1} a_{ik} \left(\prod_{j=1}^{n} P_{jt}^{a_{ij}} \right) \frac{W_t^{\alpha_i}}{P_{kt}} Y_{fit}. \tag{21}$$

Price setting. After solving for the input choice, firm (i, f)'s profit function can be expressed as follows

$$\Pi(P_{fit}, \mathbf{P}_t, Y_{it}, W_t, Z_{it}) = (1 - \tau_i) \left(\frac{P_{fit}}{P_{it}} \right)^{-\theta_i} Y_{it} - Z_{it}^{-1} \left(\prod_{k=1}^{n} P_{kt}^{a_{ik}} \right) W_t^{\alpha_i} \left(\frac{P_{fit}}{P_{it}} \right)^{-\theta_i} Y_{it}, \tag{22}$$

where $\mathbf{P}_t \equiv (P_{it})_{i \in I}$ is the vector of sectoral prices.

To circumvent the complexity of characterizing analytical rational inattention models, we adopt the usual approach of taking a second-order approximation to the firms' information acquisition and price-setting problems around the full-information equilibrium.[3]

[3] The full-information equilibrium is efficient since prices are fully flexible and the government's industry-specific taxes undo monopolistic distortions.

The approximated firms' profit function is

$$\pi(p_{fit}, \mathbf{p}_t, y_{it}, w_t, z_{it}) = \Pi_1 p_{fit} + \frac{\Pi_{11}}{2} p_{fit}^2 + \Pi_{12}' \mathbf{p} p_{fit} + \Pi_{13} y_i p_{fit} + \Pi_{14} w p_{fit} + \Pi_{15} z_i p_{fit} + \text{terms independent of } p_{fit},$$

where all the derivatives of the original profit function are evaluated at the full-information equilibrium. A small letter denotes the log-deviation of the corresponding variable from its steady-state value.

After computing the derivatives of the original profit function at the full-information equilibrium, it follows that a firm's optimal price p_{fit}^* is equal to its industry-specific marginal cost:

$$p_{fit}^* = \alpha_i w_t + \sum_{k=1}^{n} a_{ik} p_{kt} - z_{it} \equiv mc_{it}. \tag{23}$$

Since all firms within an industry have the same marginal cost (or the same optimal price), hereafter, we will drop the f subscript and denote it by p_{it}^*.

Information acquisition. Following the second-order approximation and using the solution to firms' input choice, along with the equilibrium wage condition $w = \delta_w m$, the firms' problem (16) becomes

$$\begin{aligned} &\min_{\kappa_{fit}, S_{fit}, p_{fit}(S_{fit})} \frac{1}{2} B_i \mathbb{E}_{fit}\left[(p_{fit}(S_{fit}) - p_{it}^*)^2\right] + \omega \kappa_{fit} \\ &\text{s.t.}: \quad p_{it}^* = \alpha_i \delta_w m_t + \sum_{k=1}^{n} a_{ik} p_{kt} - z_{it}, \end{aligned} \tag{24}$$

where $B_i \equiv |\Pi_{11}| = \lambda_i \theta_i$ is the curvature of the profit function around the full-information equilibrium, which equals the product of the industry-specific elasticity of substitution across varieties, θ_i, and sector i's Domar weight at the full-information equilibrium, λ_i.

As in single-agent rational inattention settings (Maćkowiak, Matějka, and Wiederholt 2018; Afrouzi and Yang 2021), it can be shown that in all equilibria, each firm optimally acquires a single signal collinear with its price. We refer to strategies with this property as *recommendation strategies*: $p_{if} = S_{if}, \;\; S_{if} \in \mathbb{S}$. The intuition is that all alternative strategies are weakly dominated by some feasible recommendation strategy.

Equilibrium prices are given by:

$$p_{fit} = \mu_{fit}\left(\alpha_i \delta_w m_t + \sum_{k=1}^{n} a_{ik} p_{kt} - z_{it}\right) + \nu_{fit}, \quad \nu_{fit} \perp \left(m_t, z_{it}, (S_{i'f't})_{(i',f') \neq (i,f)}\right). \tag{25}$$

In the equation above, $\mu_{fit} \equiv 1 - e^{-2\kappa_{fit}}$ represents the Kalman gain on the signal, where κ_{fit} is the attention capacity. We call μ_{fit} the firm's attention level (to its own marginal cost). The term ν_{fit} represents the noise in prices induced by rational inattention and has the following stochastic properties:

$$\mathbb{E}[\nu_{fit}] = 0, \quad Var[\nu_{fit}] = \mu_{fit}(1 - \mu_{fit}) Var[p^*_{it}]. \tag{26}$$

A proof for the characterization of equilibrium prices above, as well as the feasibility and optimality of recommendation strategies, is provided in Afrouzi (2024).

The information acquisition problem (24) is identical among all firms within the same industry and across time, due to the stationarity of the environment. This implies $\mu_{fit} = \mu_i$ for all firms $f \in [0, 1]$ in each industry $i \in I$ and for every period $t \in \{0, 1, \ldots\}$.

Attention capacity. As we have solved firms' signal selection problem for a given attention capacity κ_{if}, we can rewrite firms' information problem just in terms of κ_{if} as follows:

$$\min_{\kappa_i \geq 0} \left\{ \frac{1}{2} e^{-2\kappa_i} \theta_i \lambda_i \mu_i V^*_i + \omega \kappa_i \right\},$$

where $V^*_i \equiv Var(p^*_{it}) = Var\left(\alpha_i \delta_w m_t + \sum_{k=1}^{n} a_{ik} p_{kt} - z_{it}\right)$ is the variance of firm (i, f)'s optimal price, or of its industry-specific marginal cost. Since this problem is identical to all firms in the same industry, we drop the f subscript. The solution to this problem is:

$$\kappa_i = \frac{1}{2} \max \left\{ 0, \ln\left(\frac{\theta_i \lambda_i \mu_i V^*_i}{\omega} \right) \right\}.$$

We focus on the case where all firms in the economy acquire some positive amount of information, i.e. $\kappa_i > 0$ for all $i \in I$. A sufficient condition is $\omega < \theta_i \lambda_i \alpha_i^2 \delta_w^2 \sigma_m^2 + \sigma_{z,i}^2$ for all $i \in I$, i.e., the cost of information is low enough for all firms operating in the various industries. Then, industry i's equilibrium attention is given by

$$\mu_i = 1 - \frac{\omega}{\theta_i \lambda_i V^*_i}, \tag{27}$$

The equation above captures the trade-off firms face when choosing the optimal attention level. They balance the benefits of more precise information against the cost of attention, ω. The benefits depend on two industry-specific components. The first, $B_i = \theta_i \lambda_i$, reflects the curvature of the profit function with respect to pricing mistakes. Firms in industries with a higher elasticity of substitution (θ_i) face steeper competitive losses from mispricing, as well as when the industry Domar weight (λ_i) is higher. The second component, V_i^*, is the variance of marginal costs. Industries facing more volatile marginal costs – whether due to aggregate demand exposure, upstream price transmission, or idiosyncratic productivity shocks – have more to gain from acquiring information. Together, these forces imply that firms in industries where pricing mistakes are costly and marginal costs are volatile will optimally choose higher attention levels.

4.2. Equilibrium characterization

Rewriting conditions (25) and (27) in terms of industry outcomes in matrix and vector forms, we can characterize the equilibrium as follows:

PROPOSITION 1. *In equilibrium, sectoral prices and attention levels are given by*

$$\boldsymbol{p}_t = \mathrm{M}(I - A\mathrm{M})^{-1}(\boldsymbol{\alpha}\delta_w m_t - \boldsymbol{z}_t) \tag{28}$$

$$\mathrm{M} = I - \omega\, diag(\boldsymbol{\theta\lambda}')^{-1}\left(Var(\mathrm{M}^{-1}\boldsymbol{p}_t) \odot I\right)^{-1}, \tag{29}$$

where $\boldsymbol{p} \equiv (p_i)_{i \in I}$, $\mathrm{M} \equiv diag(\boldsymbol{\mu})$ and $\odot$ denotes the Hadamard product (element-wise matrix multiplication).

PROOF. See Appendix A.1. □

The equilibrium is characterized by a fixed point between price responses and attention levels. Intuitively, lower attention in a specific industry leads to dampened price responses to aggregate and sectoral shocks, which decreases the volatility of marginal costs in other sectors and encourages further inattention within those sectors. This is a highly non-linear system of equations, with no closed-form solution for a general input-output structure $(A, \boldsymbol{\alpha})$.[4] Still, we derive several insightful results regarding the determinants and macroeconomic consequences of heterogeneous firms' expectations in general production networks, which we explore in the next section.

[4] We refer the reader to Fang et al. (2024) for details about the proof of existence and uniqueness in models with endogenous inattention in high-dimensional settings.

5. Heterogeneous Firms' Expectations: Causes and Consequences

In this section, we examine the determinants and macroeconomic implications of firms' heterogeneous attention. On the determinants side, we show that inattention propagates downstream through the production network and derive a decomposition of firms' CPI expectations that connects forecast accuracy to industry characteristics. On the implications side, we show that heterogeneous attention alters the aggregate transmission of shocks and revisit the standard result that input-output linkages flatten the Phillips curve (Basu 1995; Rubbo 2023), showing that it need not hold when attention is endogenous to the production structure.

5.1. Determinants of firms' inflation expectations

Proposition 2 decomposes the industry-average firms' expectations about the consumer price index (CPI) inflation rate into three components: change in industry-specific marginal costs, attention to industry-specific marginal costs, and comovement between industry-specific marginal cost and aggregate CPI.

PROPOSITION 2. *Industry-average CPI inflation forecasts are given by*

$$\bar{\mathbb{E}}_i[\pi_t] = \underbrace{\frac{Cov(p^*_{it}, p^{CPI}_t)}{Var(p^*_{it})}}_{\text{comovement between mc and CPI}} \quad \underbrace{\mu_i}_{\text{attention to own mc}} \quad \underbrace{\Delta p^*_{it}}_{\text{realized mc}}, \tag{30}$$

where $p^{CPI}_t \equiv \boldsymbol{\gamma}' \boldsymbol{p}_t$ *is the log CPI and* $\pi_t \equiv p^{CPI}_t - p^{CPI}_{t-1}$ *is the CPI inflation rate.*

PROOF. See appendix A.3. □

The intuition behind the result rests on the fact that rationally inattentive firms acquire information about their industry-specific marginal costs, which is the key variable for their profit maximization problem. Thus, firms have more precise inflation expectations when their marginal costs exhibit a stronger comovement with the price index. Moreover, industries optimally choose different attention levels, due to heterogeneous incentives to acquire information. Hence, the inflation expectations of firms in more attentive industries are more responsive to changes in their industry-specific marginal cost. At the firm level, inflation forecasts can be decomposed into two components: an industry-average component detailed in (30) and an idiosyncratic noise term.

The decomposition in Proposition 2 yields a direct connection between the model's primitives and the accuracy of firms' inflation forecasts. Specifically, by equation (30), it follows that the expected squared forecast error for a firm in industry i is

$$\mathbb{E}\left[FE^2_{fit}[\pi_t]\right] = \left(1 - \mu_i R_i^2\right) Var(\pi_t), \tag{31}$$

where R_i^2 is the R-squared from a regression of CPI inflation π_t on changes in sector i's prices ($\Delta\pi_{it}$) or marginal costs (Δmc_{it}).

Equation (31) decomposes forecast accuracy into two industry-specific components: the attention level μ_i and the degree of comovement between the industry's marginal cost and the aggregate price level R_i^2. What matters for forecast accuracy is the *product* $\mu_i R_i^2$: a highly attentive firm still produces inaccurate CPI forecasts if its marginal cost does not comove with aggregate inflation (R_i^2 low), and a firm in an industry highly representative of aggregate conditions still makes large forecast errors if it is not paying attention (μ_i low). Accurate inflation forecasts thus require both that the firm acquires precise information *and* that this information is relevant for aggregate price movements.

This decomposition provides a direct link between the model and the empirical findings documented in Section 2.3. In particular, the model predicts that firms in high-Domar-weight industries make smaller forecasting mistakes because their marginal costs are more representative of aggregate price movements, resulting in a higher R_i^2, holding attention levels constant. Similarly, firms in industries with more flexible prices make more accurate forecasts because price flexibility maps directly into higher attention levels μ_i in the model.

Finally, it is worth noting that equation (31) provides a way to recover industry-level attention μ_i from data on firms' expectations.

5.2. Downstream compounding of inattention and price rigidities

Our first result shows how inattention, the source of price rigidities in our model, in a particular sector affects attention levels in other sectors of the economy.

LEMMA 1. *Inattention compounds downstream. Concretely, holding other sectors' attention levels fixed, sector i's optimal attention is weakly increasing in sector j's attention level:*

$$\frac{\partial \mu_i^{BR}(\mu)}{\partial \mu_j} \geq 0,$$

where $\mu_i^{BR}(\boldsymbol{\mu})$ *denotes sector i's best-response attention level for any sector* $i \in I$ *given a strategy profile* $\boldsymbol{\mu} \in \mathbb{R}^n$.

PROOF. See appendix A.2. □

This result shows that inattention compounds downstream in the production network. The (i, j)-th element of the matrix $(I - AM)^{-1}$ corresponds to the elasticity of sector i's marginal cost with respect to sector j's marginal cost, directly and indirectly through the input-output network. Therefore, the optimal attention choice by industry j will influence industry i's information acquisition decision as long as those industries are connected through the network structure of the economy. The economic mechanism is as follows: when upstream firms are less attentive, their prices respond less to shocks, making downstream firms' input costs—and hence marginal costs—less volatile. This lower marginal cost volatility decreases the incentives for information acquisition for downstream firms, inducing them to also become less attentive.

5.3. Heterogeneous attention and the aggregate effects of shocks

Having established how industry characteristics determine firms' attention and forecast accuracy, we now investigate the aggregate consequences: whether heterogeneous attention amplifies or dampens the transmission of shocks to nominal and real variables.

LEMMA 2. *Let* $p^{\phi} \equiv \phi' \boldsymbol{p}$ *be an aggregate price index for any weight* ϕ *in the simplex* Δ^n, *and* $\tilde{\boldsymbol{\delta}}^x \equiv d\boldsymbol{p}^*/dx$ *be the vector of marginal costs responses to a shock* $x \in \{m, \boldsymbol{z}\}$. *The effect of a shock to x on the aggregate price level is given by*

$$\frac{dp_t^{\phi}}{dx_t} = \underbrace{\mathbb{E}_{\phi}[\mu]\mathbb{E}_{\phi}\left[\tilde{\delta}^x\right]}_{\text{representative agent channel}} + \underbrace{Cov_{\phi}(\mu, \tilde{\delta}^x)}_{\text{heterogeneous attention channel}} .$$

PROOF. See Appendix A.4. □

The first term captures the representative agent channel: the average attention and the average response of marginal costs to the shock shape its propagation. The second term captures the insight that the aggregate effect of shocks will be larger if the more attentive sectors are the ones whose marginal costs respond more to the shock. For instance, there is more monetary non-neutrality if marginal costs are more responsive to changes in aggregate nominal demand in less attentive industries.

The sign of the covariance term depends on all the primitives and does not have an analytical representation. The covariance is positive (negative) —so that heterogeneous attention amplifies the aggregate price response—when industries whose marginal costs are more exposed to the shock are also the more (less) attentive ones. Our calibrated economy exhibits a positive covariance for monetary shocks, implying that heterogeneous attention amplifies inflation responses relative to a homogeneous attention benchmark. We quantify this mechanism in Section 6.

5.4. The slope of the Phillips curve

We now turn to the implications of heterogeneous attention for the Phillips curve. In general, the dynamic model described in Section 3 does not admit closed-form solutions for firms' equilibrium signals and prices; we therefore derive the Phillips curve for the special case of myopic firms ($\beta = 0$) with persistent monetary shocks ($\rho_m > 0$) and study the determinants of its slope.

PROPOSITION 3. *If $\beta = 0$ and $\rho_m > 0$, sectoral Phillips curves are*

$$\boldsymbol{\pi}_t = \mathcal{B} y_t + (I - \mathcal{V})\bar{\mathbb{E}}_{t-1}[\Delta \boldsymbol{mc}_t] - \mathcal{V}\left(p_{t-1} + (I-A)^{-1}\boldsymbol{z}_t\right), \tag{32}$$

where $\Delta \boldsymbol{mc}_t \equiv mc_t - mc_{t-1}$ and

$$\begin{aligned} \mathcal{B} &\equiv \frac{\mathrm{M}(I - A\mathrm{M})^{-1}\boldsymbol{\alpha}}{1 - \boldsymbol{\gamma}'\mathrm{M}(I - A\mathrm{M})\boldsymbol{\alpha}}, \\ \mathcal{V} &\equiv [\mathrm{M}(I - A\mathrm{M})^{-1} - \mathcal{B}((I-A)^{-1}\boldsymbol{\gamma} - \boldsymbol{\gamma}'\mathrm{M}(I - A\mathrm{M})^{-1})](I - A). \end{aligned} \tag{33}$$

PROOF. See Appendix A.5. □

The Phillips curve in (32) resembles its counterpart in the full-information production network model of Rubbo (2023), with one key difference: the attention matrix M, which governs nominal rigidities, is endogenous to the economy's input-output structure rather than exogenously imposed. Note that sectoral inflation rates depend on sectors' expectations about their own marginal costs, not aggregate inflation. Inattention M and network A interact in a nonlinear way to shape Phillips curves, and the expectations term appears in (32) due to the anchoring of expectations to prior beliefs induced by the Kalman filter.

Our framework provides a qualification for an important result of the literature on inflation and networks. Several papers examining price setting in production networks (Basu 1995; Rubbo 2023, among others,) find that Phillips curves are flatter in economies with larger intermediate input shares, as network linkages compound nominal frictions present in each industry. In our setting, the effects of changes in the input-output linkages on the Phillips curve slope are ambiguous, as equilibrium attention levels are endogenous to the input-output structure. Corollary 1 formalizes this insight.

COROLLARY 1. *The slope of sectoral Phillips curves, $\mathcal{B}$, is weakly increasing in each attention level $(\mu_i)_{i \in I}$. However, the effects of changes in the input-output structure of the economy $(A, \boldsymbol{\alpha})$ on $\mathcal{B}$ are ambiguous.*

PROOF. See Appendix A.6. □

Intuitively, stronger input-output linkages have two opposing effects. On one hand, they compound nominal frictions across sectors, flattening the Phillips curve as in the standard framework. On the other hand, stronger linkages raise the volatility of marginal costs, which increases optimal attention and thereby steepens the Phillips curve. In Section 6, we show that the second force partially offsets the first: the endogenous adjustment of information acquisition to the production structure mutes the flattening effect that stronger input-output linkages would have under exogenous nominal rigidities.

6. Quantitative Results

The analytical results in the previous section established that heterogeneous attention propagates downstream, shapes firms' inflation forecasts, and alters the aggregate transmission of shocks. We now return to the general model with forward-looking decisions and persistent shocks ($\beta, \rho_m > 0$) outlined in Section 3 to quantify the macroeconomic implications of the interplay between endogenous inattention and production networks.

6.1. Calibration

Table 3 summarizes our baseline calibration. The model is calibrated to US quarterly data.

We calibrate the input-output parameters—intermediate input shares A, labor shares $\boldsymbol{\alpha}$, and final consumption shares $\boldsymbol{\gamma}$—using the 2022 Bureau of Economic Analysis (BEA) input-output tables. Industries are aggregated using Domar weights to match the 15-industry classification in the SoFIE data.

We estimate the persistence and volatility of nominal GDP growth directly from US data, obtaining $\rho_m = 0.265$ and $\sigma_m = 0.011$. The discount factor $\beta = 0.98^{1/4}$ implies an annual real interest rate of 2 percent. The covariance matrix of industry-specific TFP shocks, Σ,

TABLE 3. Calibrated parameters

Parameter	Description	Value	Source
α	Labor shares		2022 BEA IO table
A	Input expenditure shares		2022 BEA IO table
γ	Consumption shares		2022 BEA IO table
ω	Information cost	0.023	Anchoring to prior (34)
β	Discount factor	$0.98^{1/4}$	annual real rate = 2%
ρ_m	Persistence of NGDP growth	0.265	US NGDP
σ_m	Std of NGDP growth	0.011	US NGDP
δ_w	Wage rigidity	0.73	$Cov(\Delta m, \Delta w)$
Σ	Cov. matrix of sectoral TFP		ILPA/KLEMS
θ	Industry CES	6	20% markup

is estimated using data from the BEA/BLS Integrated Industry-Level Production Account (ILPA) and KLEMS databases.

We set the wage rigidity parameter $\delta_w = 0.73$ to match the covariance between nominal GDP growth and wage growth in US data. The elasticity of substitution across varieties within each industry is set to $\theta = 6$, implying a steady-state markup of 20 percent.

We calibrate the information cost parameter $\omega = 0.023$ to match the degree to which firms anchor their inflation forecasts to their prior beliefs, following Afrouzi (2024) and Maćkowiak and Wiederholt (2015). Specifically, we use the following regression from the Survey of Firms' Inflation Expectations (SoFIE) data:

$$\mathbb{E}_{fit}[\pi_t] = \chi_t + \Gamma \mathbb{E}_{fit-4}[\pi_t] + \text{error}_{it} \tag{34}$$

The intuition behind this calibration is that with a higher information cost ω, firms' signals are less accurate, and hence firms rely more on their priors when forming beliefs. We estimate an anchoring coefficient of $\Gamma = 0.242$. See Appendix C.3 for details on this estimation.

6.2. Solution method

We solve the model using an iterative fixed-point algorithm that adapts the method of Afrouzi and Yang (2021) to a multi-industry setting; a detailed description is provided in Appendix B.

The key computational challenge is the nonlinear fixed point between sectoral prices and attention levels: prices depend on attention through the degree of nominal rigidity, while optimal attention depends on prices through the volatility of marginal costs. We proceed as follows. Given a guess for the joint Gaussian process of sectoral prices, nominal demand, and industry-specific productivities, we derive each sector's optimal pricing strategy in a symmetric stationary equilibrium. This determines the stochastic process for sectoral marginal costs, which we approximate as an integrated moving average in monetary and productivity shocks. We then solve each sector's rational inattention problem using the DRIP method of Afrouzi and Yang (2021), which is efficient enough to permit the many iterations required to calibrate ω. The resulting beliefs and pricing strategies yield an updated joint process for sectoral prices, and we iterate until convergence.[5]

6.3. Model validation

We assess the model's ability to replicate the empirical relationship between industry characteristics and forecast accuracy documented in section 2.3. Specifically, we simulate time series from the calibrated model and regress industry-level average absolute nowcast errors on Domar weights and price flexibility, mirroring the empirical exercise.

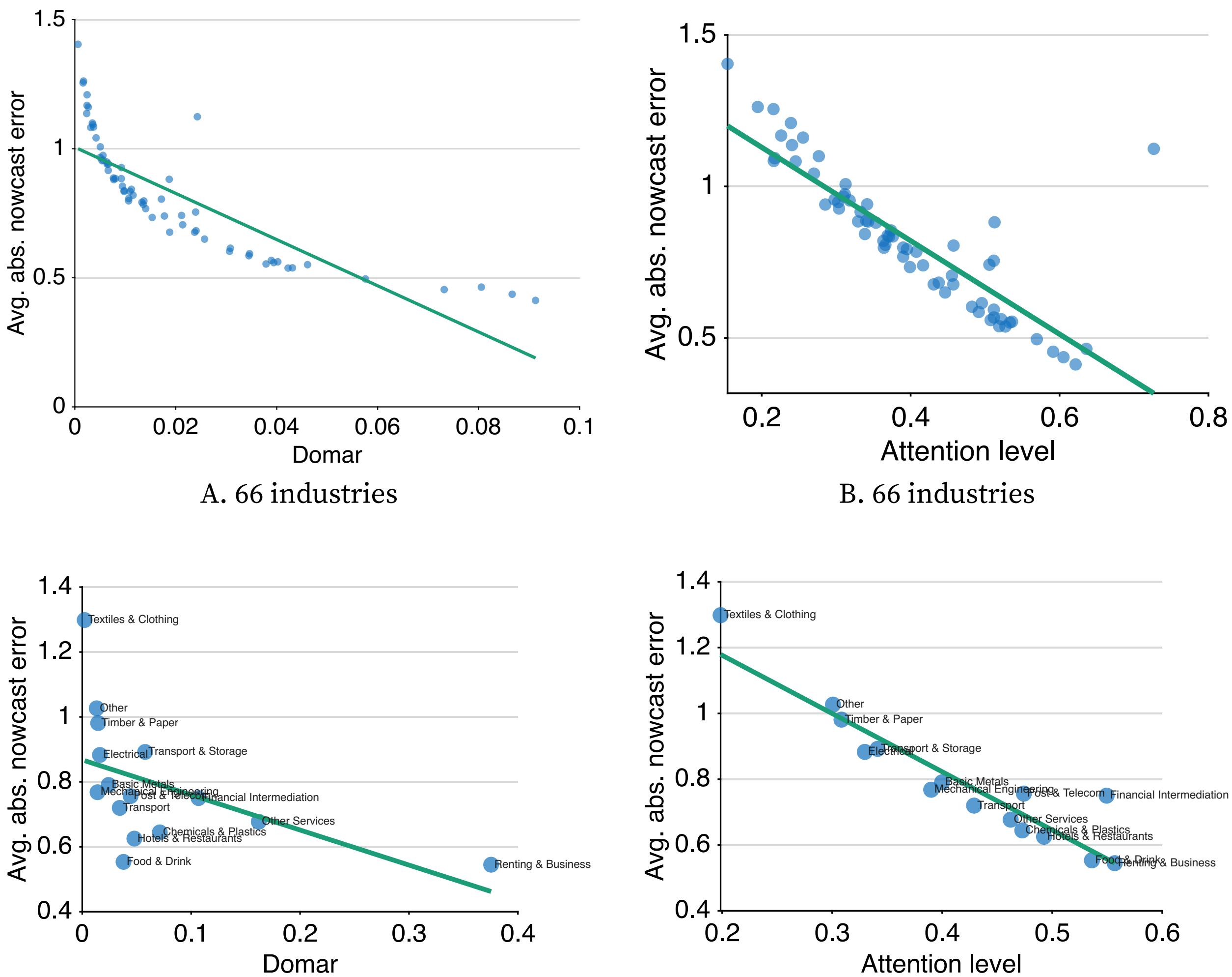
Figure 2 presents scatter plots of the model-implied relationship between the industry-average absolute forecast errors and industry characteristics. The left panel shows the relationship between Domar weights and forecast accuracy. Industries with larger Domar weights tend to have smaller forecast errors, consistent with our empirical findings. In the model, industries with higher Domar weights are more central in the production network and thus their marginal costs comove more strongly with aggregate prices. Therefore, they can infer aggregate inflation more accurately given their noisy observation of their industry-specific marginal costs. The right panel displays the relationship between attention levels and price flexibility. Our model captures the positive correlation between price flexibility and forecast accuracy, as firms that adjust prices more frequently have stronger incentives to pay attention to inflation.

Testing the mechanism. The theory presented in Sections 3 and 5 makes a sharp prediction regarding the firms' belief formation process about the price level, given by:

$$\mathbb{E}_{fit}[p^{CPI}] = \mu_i \frac{Cov(p_i^*, p^{CPI})}{Var(p_i^*)} p_i^* + u_{fit}, \quad u_{fit} \perp p_{it}^*. \tag{35}$$

[5] As pointed out by Afrouzi (2024), attention capacity is always strictly positive in the dynamic model, unlike in the static case. The unit root in m_t implies that under a zero-capacity strategy, uncertainty about marginal costs grows without bound, so firms always find it optimal to acquire some information.

FIGURE 2. Industry characteristics and expectation accuracy in the model



A. 66 industries

B. 66 industries

C. 15 industries

D. 15 industries

Note: The figure shows scatter plots of the model-implied relationship between industry-average absolute nowcast errors of inflation and industry characteristics. Each point represents an industry, with the x-axis showing the industry characteristic and the y-axis showing the average absolute nowcast error from simulated data. The top row uses the 66-industry classification from the BEA input-output tables. The bottom row aggregates to the 15-industry classification used in the SoFIE data by computing sales-weighted averages of industry characteristics and forecast errors within each SoFIE industry grouping.

The equation above yields two predictions. First, firms' inflation expectations should strongly comove with their marginal costs. Second, the coefficient should vary across industries according to differences in their attention levels and the comovement between marginal costs and the CPI.

6.4. Propagation of monetary shocks

We now use the calibrated model to quantify the three aggregate implications established in Section 5: how production networks shape the transmission of monetary shocks, whether the network effect operates primarily through real or nominal rigidities, and whether heterogeneous attention matters relative to the homogeneous benchmark commonly assumed in the literature.

Figure 3 plots the response of aggregate output and inflation to a one-percentage-point monetary shock in the baseline economy and in an economy without input-output linkages (a 'horizontal' economy).

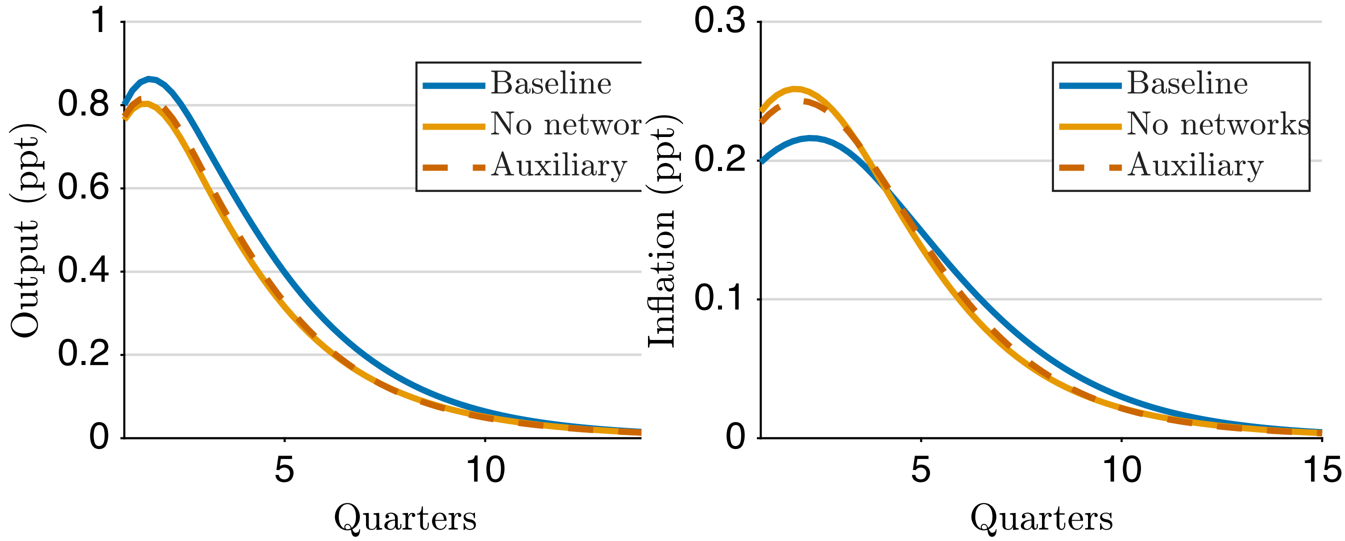
The left panel shows that the output response to a monetary shock is larger and more persistent in the baseline economy with production networks compared to an economy without input-output linkages (a 'horizontal' economy). The right panel reveals that production networks dampen the inflation response: the initial inflation increase is around 0.20 percentage points in the baseline economy compared to 0.24 percentage points without networks. Production networks thus amplify the real effects of monetary policy while attenuating its nominal effects—consistent with the standard intuition that input-output linkages compound nominal frictions across sectors, slowing the pass-through of demand shocks to prices and prolonging their effect on output.

A natural question is whether this operates primarily through real rigidities—the mechanical compounding of input costs through the network—or through nominal rigidities, which in our model reflect firms' endogenous information acquisition decisions. Figure 3 decomposes the role of these two channels by introducing an auxiliary economy. The auxiliary economy shuts down production networks but holds firms' attention levels fixed at those from the baseline economy, thereby isolating the contribution of real rigidities from input-output linkages.

The auxiliary economy's impulse responses closely track those of the horizontal economy in both panels: the output and inflation responses are nearly identical. In contrast, the baseline economy—which allows attention to adjust endogenously to the presence of networks—exhibits noticeably different dynamics. This comparison reveals that most of the difference between the baseline and no-networks economies stems from real rigidities embedded in the production structure, while differences in endogenous information acquisition play a secondary role.

It is important to interpret this decomposition correctly. The auxiliary economy holds attention levels *fixed* at their baseline values while removing input-output linkages, thereby isolating the contribution of real rigidities from that of endogenous attention. The fact that

FIGURE 3. Impulse responses to a monetary shock: role of networks



Note: Impulse responses of output (left) and CPI inflation (right) to a one-percentage-point monetary shock. The baseline economy features production networks with endogenous attention. The horizontal economy has no input-output linkages and endogenous attention. The auxiliary economy has no input-output linkages but retains the same attention levels as the baseline economy.

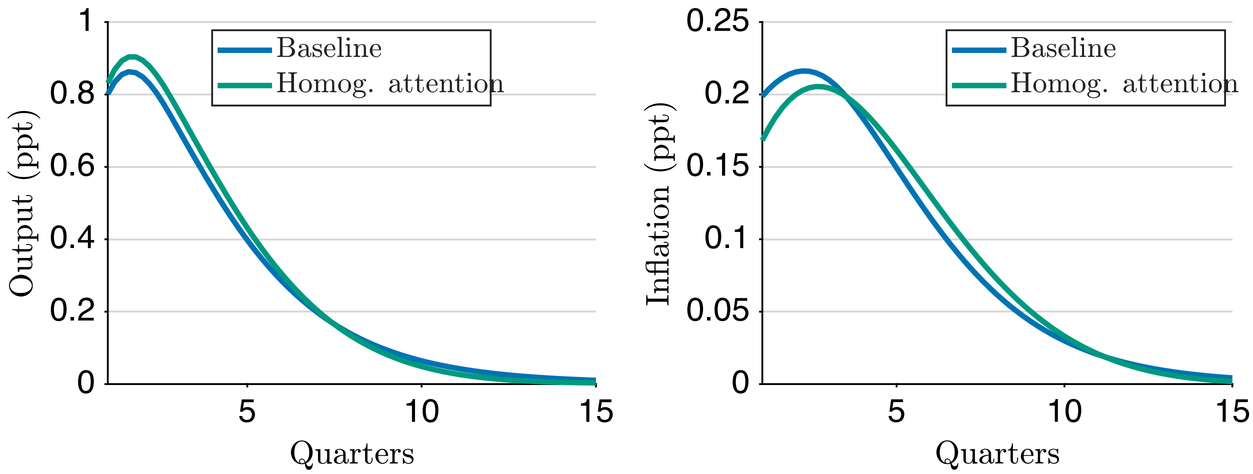
real rigidities are the dominant channel does not diminish the role of heterogeneous attention in our framework: the key contribution of endogenous attention operates not through the *average level* of information acquisition, but through its *cross-sectional distribution* across industries. As we show next, the pattern of which industries are attentive—shaped by network position—has distinct aggregate implications that cannot be captured by an economy with homogeneous attention.

The role of heterogeneous attention. We now isolate the quantitative importance of heterogeneous attention across industries. Figure 4 compares the baseline economy, where attention varies endogenously across industries, with a counterfactual economy featuring homogeneous attention levels. In the counterfactual, all firms receive exogenous noisy signals about each shock $x \in \{m, \boldsymbol{z}\}$ of the form

$$S^x_{fit} = x_t + u^x_{fit}, \quad u^x_{fit} \sim \mathcal{N}(0, \tau_x \sigma^2_x), \tag{36}$$

where τ_x is a common signal-to-noise ratio calibrated so that the average information quality matches the average attention level in the baseline economy. This counterfactual corresponds to the exogenous and homogeneous incomplete information assumption commonly adopted in the literature (Angeletos and Huo 2021; La'O and Tahbaz-Salehi 2022), and thus allows us to quantify the role played by endogenous heterogeneity in attention.

FIGURE 4. Impulse responses to a monetary shock: heterogeneous vs. homogeneous attention



Note: Impulse responses of output (left) and CPI inflation (right) to a one-percentage-point monetary shock. The baseline economy features endogenous heterogeneous attention across industries. The homogeneous attention economy assigns all firms the same signal quality, calibrated to match the Domar-weighted average attention in the baseline.

The baseline economy with heterogeneous attention exhibits a stronger inflation response to monetary shocks than the homogeneous attention counterfactual. To understand this result, recall from Lemma 2 that the effect of a shock on aggregate inflation decomposes into two terms: a representative-agent channel, which depends on the average attention level and the average marginal cost exposure, and a heterogeneous attention channel, which depends on the covariance between industry attention levels and their marginal cost responses. When this covariance is positive—that is, when the industries that are more attentive are also those whose marginal costs respond more strongly to the shock—heterogeneous attention amplifies the aggregate inflation response relative to a world with homogeneous attention.

In our calibrated economy, this covariance is unambiguously positive due to a selection effect. Industries with large Domar weights have higher equilibrium attention, as the stakes from mispricing are larger for firms that contribute more to aggregate price movements. In the data, these high-Domar-weight industries also tend to be labor-intensive. Labor intensity increases the exposure of marginal costs to monetary shocks: a monetary expansion raises aggregate demand and wages, and labor-intensive industries experience larger increases in their marginal costs. The confluence of these two forces—high attention driven by Domar weights and high exposure driven by labor intensity in the same industries—generates a positive covariance between attention and marginal cost exposure.

This positive covariance steepens the aggregate Phillips curve. When high-exposure industries pay closer attention to aggregate conditions, they adjust their prices more accurately in response to monetary shocks, passing through a larger share of the shock to the price level. In contrast, when attention is equalized across industries as in the counterfactual, high-exposure sectors become relatively less attentive than in the baseline. Their price adjustments become less responsive to the true state of aggregate demand, muting the pass-through of the shock to prices. The result is a flatter Phillips curve: the homogeneous attention economy exhibits a smaller inflation response and, consequently, a larger output response to monetary shocks.

7. Conclusion

This paper investigates the causes and macroeconomic consequences of heterogeneity in firms' inflation expectations across industries, combining new empirical evidence with a tractable rational inattention model of price-setting firms operating within a production network.

On the empirical side, we combine a comprehensive dataset on US firms' inflation expectations with industry-level data on production structure and pricing behavior. We find that the accuracy of firms' inflation forecasts is systematically associated with industry attributes: firms in industries with larger Domar weights and more flexible prices have significantly more accurate inflation forecasts. We also find that firms form inflation expectations based on the realized industry-specific marginal costs, consistent with the view that firms use their own cost conditions as a signal about aggregate inflation.

On the theoretical side, two key mechanisms emerge from the model. First, inattention compounds downstream through the production network: low attention in upstream industries dampens their price responses, reducing marginal cost volatility in downstream sectors and discouraging further information acquisition. Second, firms' inflation expectations are shaped by their attention to industry-specific marginal costs, the comovement between marginal costs and the aggregate price level, and the realized marginal cost itself—linking forecast accuracy to both industry-specific incentives and the informativeness of cost conditions about aggregate inflation.

Quantitatively, the calibrated model shows that production networks attenuate the inflation response to monetary policy, while amplifying and prolonging the real output response. Which industries are attentive matters for aggregate dynamics: industries with large Domar weights are more attentive, and in the data these industries also tend to be labor-intensive and thus more exposed to monetary shocks, generating a selection effect that steepens the aggregate Phillips curve relative to a homogeneous attention benchmark. This selection effect partially offsets the standard flattening of the Phillips curve from input-output linkages.

References

Afrouzi, Hassan. 2024. "Strategic Inattention, Inflation Dynamics, and the Nonneutrality of Money." *Journal of Political Economy* 132 (10): 3378–3420.

Afrouzi, Hassan, and Choongryul Yang. 2021. "Dynamic Rational Inattention and the Phillips Curve." CESifo Working Paper 8840, CESifo Group Munich.

Andrade, Philippe, Olivier Coibion, Erwan Gautier, and Yuriy Gorodnichenko. 2022. "No firm is an island? How industry conditions shape firms' expectations." *Journal of Monetary Economics* 125: 40–56.

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

Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry. 2012. "Measuring the Upstreamness of Production and Trade Flows." *American Economic Review: Papers & Proceedings* 102 (3): 412–16.

Basu, Susanto. 1995. "Intermediate Goods and Business Cycles: Implications for Productivity and Welfare." *The American Economic Review* 85 (3): 512–531.

Blanchard, Olivier, and Jordi Galí. 2007. "Real Wage Rigidities and the New Keynesian Model." *Journal of Money, Credit and Banking* 39: 35–65.

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko. 2024. "The inflation expectations of U.S. firms: Evidence from a new survey." *Journal of Monetary Economics* 145: 103569.

Carvalho, Vasco M., and Alireza Tahbaz-Salehi. 2019. "Production Networks: A Primer." *Annual Review of Economics* 11 (Volume 11, 2019): 635–663.

Coibion, Olivier, Yuriy Gorodnichenko, and Saten Kumar. 2018. "How Do Firms Form Their Expectations? New Survey Evidence." *American Economic Review* 108 (9): 2671–2713.

Fang, Yue, Zhen Huo, Minghao Li, and Jieran Wu. 2024. "Inattentive Network: Evidence and Theory." Working paper, Zhejiang University, Yale University, Peking University.

Federal Reserve of Cleveland. 2023. "Detailed Methodology for the Survey of Firms' Inflation Expectations (SoFIE)."

Gagliardone, Luca, and Joris Tielens. 2025. "Dynamic Pricing under Information Frictions: Evidence from Firm-level Subjective Expectations." Working Paper.

Golosov, Mikhail, and Robert E. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2): 171–199.

Hajdini, Ina, Saten Kumar, Samreen Malik, Jordan J. Norris, and Mathieu Pedemonte. 2025. "Supply Chain Networks and the Macroeconomic Expectations of Firms." Working Paper 25-17, Federal Reserve Bank of Cleveland.

La'O, Jennifer, and Alireza Tahbaz-Salehi. 2022. "Optimal Monetary Policy in Production Networks." *Econometrica* 90 (3): 1295–1336.

Lucas, Robert E. 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (2):

103–124.

Mackowiak, Bartosz, and Mirko Wiederholt. 2009. “Optimal Sticky Prices under Rational Inattention.” *American Economic Review* 99 (3): 769–803.

Mankiw, N. Gregory, and Ricardo Reis. 2002. “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve*.” *The Quarterly Journal of Economics* 117 (4): 1295–1328.

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt. 2018. “Dynamic rational inattention: Analytical results.” *Journal of Economic Theory* 176: 650–692.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2015. “Business Cycle Dynamics under Rational Inattention.” *The Review of Economic Studies* 82 (4): 1502–1532.

Minton, Robert, and Hugo Monnery. 2025. “How Firms Form Beliefs and the Implications for Inflation.” Working Paper.

Nimark, Kristoffer. 2008. “Dynamic pricing and imperfect common knowledge.” *Journal of Monetary Economics* 55 (2): 365–382.

Pastén, Ernesto, Raphael Schoenle, and Michael Weber. 2024. “Sectoral Heterogeneity in Nominal Price Rigidity and the Origin of Aggregate Fluctuations.” *American Economic Journal: Macroeconomics* 16 (2): 318–52.

Rubbo, Elisa. 2023. “Networks, Phillips Curves, and Monetary Policy.” *Econometrica* 91 (4): 1417–1455.

Sims, Christopher A. 2003. “Implications of rational inattention.” *Journal of Monetary Economics* 50 (3): 665–690.

Weber, Michael, Bernardo Candia, Hassan Afrouzi, Tiziano Ropele, Rodrigo Lluberas, Serafin Frache, Brent Meyer, Saten Kumar, Yuriy Gorodnichenko, Dimitris Georgarakos, Olivier Coibion, Geoff Kenny, and Jorge Ponce. 2025. “Tell Me Something I Don’t Already Know: Learning in Low- and High-Inflation Settings.” *Econometrica* 93 (1): 229–264.

Woodford, Michael. 2002. “Imperfect Common Knowledge and the Effects of Monetary Policy.” In *Aghion, P., R. Frydman, J. Stiglitz, and M. Woodford (eds.) Knowledge, Information and Expectations in Modern Macroeconomics,*: Princeton University Press.

Appendix A. Proofs

A.1. Proof of Proposition 1

PROOF. Writing the condition determining marginal costs (23) in vector form yields:

$$\boldsymbol{p}_t^* = \boldsymbol{\alpha}\delta_w m_t + A\boldsymbol{p}_t - \boldsymbol{z}_t \tag{A1}$$

The equilibrium condition for industry prices in vector form is $\boldsymbol{p}_t = \mathrm{M}\boldsymbol{p}_t^*$ by (25). Substituting that in the marginal cost condition above, it follows that

$$\boldsymbol{p}_t = \mathrm{M}(I - A\mathrm{M})^{-1}(\boldsymbol{\alpha}\delta_w m_t - \boldsymbol{z}_t). \tag{A2}$$

Writing the optimal attention condition (27) in matrix form yields

$$\mathrm{M} = I - \omega \mathrm{diag}(\theta)^{-1}(\mathrm{Var}(\mathrm{M}^{-1}\boldsymbol{p}_t) \odot I)^{-1}. \tag{A3}$$

□

A.2. Proof of Lemma 1

PROOF. The covariance matrix of industry-specific marginal costs is given by:

$$Var(\boldsymbol{p}_t^*) = \tilde{L}\Sigma_0\tilde{L}', \tag{A4}$$

where $\tilde{L} \equiv (I - A\mathrm{M})^{-1}$ and $\Sigma_0 \equiv \boldsymbol{\alpha}\boldsymbol{\alpha}'\delta_w^2\sigma_m^2 + \Sigma$. We proceed in three steps.

Step 1. The entries of $\tilde{L}$ are increasing in attention levels. By the Neumann series expansion, $\tilde{L} = \sum_{k=0}^{\infty}(A\mathrm{M})^k$. Since A has non-negative entries and M is a diagonal matrix with non-negative entries, each term $(A\mathrm{M})^k$ has non-negative entries that are weakly increasing in the diagonal elements of M. Therefore, if $\tilde{L}_{ij} > 0$, then $\tilde{L}_{ij}$ is strictly increasing in μ_j.

Step 2. The variance of sector i's marginal cost is increasing in μ_j whenever $\tilde{L}_{ij} > 0$. The diagonal element $V_i^* = [\tilde{L}\Sigma_0\tilde{L}']_{ii} = \sum_{k,l}\tilde{L}_{ik}(\Sigma_0)_{kl}\tilde{L}_{il}$. Since Σ_0 is element-wise non-negative (as $\boldsymbol{\alpha}\boldsymbol{\alpha}'$ and the diagonal matrix Σ both have non-negative entries) and $\tilde{L}$ has non-negative entries, each term in the sum is non-negative. By Step 1, increasing μ_j increases $\tilde{L}_{ij}$, which strictly increases V_i^*.

Step 3. Sector i's optimal attention is increasing in V_i^*. By equation (27), $\mu_i^{BR} = 1 - \omega/(\theta_i V_i^*)$, which is strictly increasing in V_i^*.

Combining Steps 1–3, $\partial\mu_i^{BR}/\partial\mu_j > 0$ whenever $(I - AM)^{-1}_{ij} > 0$. □

A.3. Proof of Proposition 2

PROOF. Since the equilibrium is a recommendation strategy, each firm's price equals its signal at time t. Projecting the CPI onto each firm's price, it follows that

$$p_t^{CPI} = \zeta_{if} p_{fit} + u_{fit}.$$

where u_{fit} is a residual term orthogonal to p_{fit} and $\zeta_{if} \equiv Cov(p_{fit}, p_t^{CPI})/Var(p_{fit})$. Using the equilibrium price condition (25), it follows that

$$\begin{aligned}\zeta_{if} &= \frac{Cov(\mu_i p_{it}^* + \nu_{fit}, p_t^{CPI})}{Var(\mu_i p_{it}^* + \nu_{fit})} \\ &= \frac{Cov(p_{it}^*, p_t^{CPI})}{Var(p_{it}^*)} \equiv \zeta_i^*,\end{aligned}$$

where we used the fact that $\mathbb{E}[\nu_{fit}] = 0$, $Cov(\nu_{fit}, p_{it}^*) = 0$, and $Var(\nu_{fit}) = \mu_i(1-\mu_i)Var(p_i^*)$. The last equality follows from the signal structure: since $p_{fit} = \mu_i p_{it}^* + \nu_{fit}$ with ν_{fit} orthogonal to p_{it}^*, we have $Var(p_{fit}) = \mu_i^2 Var(p_{it}^*) + Var(\nu_{fit})$. Under the optimal signal structure, $Var(p_{fit}) = \mu_i Var(p_{it}^*)$, which implies $Var(\nu_{fit}) = \mu_i(1-\mu_i)Var(p_{it}^*)$. Therefore, the firm-level CPI forecast can be expressed as:

$$\mathbb{E}_{fit}[p_t^{CPI}] = \zeta_{if} p_{fit} = \frac{Cov(p_{it}^*, p_t^{CPI})}{Var(p_{it}^*)}\mu_i p_{it}^* + \frac{Cov(p_{it}^*, p_t^{CPI})}{Var(p_{it}^*)}\nu_{fit},$$

where ν_{fit} is a term orthogonal to p_{it}^*. To obtain the industry-average CPI forecast, we take the expectation over all firms within industry i. Since ν_{fit} is orthogonal to p_{it}^* and has zero mean, it drops out when taking expectations:

$$\bar{\mathbb{E}}_{it}[p_t^{CPI}] = \frac{Cov(p_{it}^*, p_t^{CPI})}{Var(p_{it}^*)}\mu_i p_{it}^*.$$

□

A.4. Proof of Lemma 2

PROOF. The effect of a shock $x \in \{m, \boldsymbol{z}\}$ on CPI is given by

$$\frac{dp_t^{CPI}}{dx_t} = \gamma' \mathrm{M} \frac{d\boldsymbol{p}_t^*}{dx_t} = \mathbb{E}_\gamma\left[\mu \frac{d\boldsymbol{p}_t^*}{dx_t}\right] = \mathbb{E}_\gamma[\mu] \times \mathbb{E}_\gamma\left[\frac{d\boldsymbol{p}_t^*}{dx_t}\right] + \mathrm{Cov}_\gamma\left(\mu, \frac{d\boldsymbol{p}_t^*}{dx_t}\right).$$

□

A.5. Proof of the Phillips curve Proposition

PROOF. Combining equilibrium prices (25) with the fact that the output gap is given by $\tilde{y}_t = m_t - \boldsymbol{\gamma}' \boldsymbol{p}_t$ yields:

$$(I - \mathrm{M}(I - A\mathrm{M})^{-1} \boldsymbol{\alpha} \boldsymbol{\gamma}' \delta_w) \boldsymbol{p}_t = \mathrm{M}(I - A\mathrm{M})^{-1}(\boldsymbol{\alpha} \delta_w \tilde{y}_t - \boldsymbol{z}_t)$$

Inverting the coefficient on inflation on the left-hand side yields the result (33). □

A.6. Proof of Corollary 1

PROOF. By the Neumann series expansion, $(I - A\mathrm{M})^{-1} = \sum_{k=0}^{\infty} (A\mathrm{M})^k$, which is increasing element-wise in attention levels $\boldsymbol{\mu}$ since each term $(A\mathrm{M})^k$ has non-negative entries that are weakly increasing in $\boldsymbol{\mu}$. Since the slope matrix $\mathcal{B}$ in (33) depends on $(I - A\mathrm{M})^{-1}$ through a composition of positive linear operations, it follows that higher attention levels increase the slope of the sectoral Phillips curves. □

Appendix B. Solution method

This section describes the numerical algorithm used to solve for the equilibrium of the dynamic model. The solution is a fixed point between the stochastic process for sectoral prices and firms' optimal attention allocation.

B.1. Solution representation

The equilibrium solution takes the form of impulse response functions of sectoral prices to monetary and productivity shocks:

$$\boldsymbol{p}_t = \Phi_m(L) \varepsilon_{mt} + \Phi_z(L) \boldsymbol{z}_t, \tag{A5}$$

where $\Phi_m(L)$ and $\Phi_z(L)$ are lag polynomials capturing the dynamic response of prices to shocks.

Since the monetary policy shock m_t follows a unit root process, we work with the cumulated monetary shock $\tilde{\varepsilon}_{mt} \equiv \sum_{s=0}^{\infty} \varepsilon_{m,t-s} = (1-L)^{-1}\varepsilon_{mt}$. Truncating to a finite lag length $\bar{L}$, the price vector can be approximated as:

$$\boldsymbol{p}_t \approx \Psi x_t, \tag{A6}$$

where Ψ is an $n \times (n+1)\bar{L}$ matrix and x_t is the state vector:

$$x_t \equiv \left(\tilde{\varepsilon}_{mt}, \tilde{\varepsilon}_{m,t-1}, \ldots, \tilde{\varepsilon}_{m,t-\bar{L}+1}, z_{1,t}, \ldots, z_{1,t-\bar{L}+1}, \ldots, z_{n,t}, \ldots, z_{n,t-\bar{L}+1}\right)'.$$

B.2. Algorithm

The algorithm iterates between solving firms' rational inattention problems and updating the equilibrium price process:

a. **Initialize**: Start with an initial guess $\Psi^{(0)}$ for the price coefficients.
b. **Solve rational inattention problems**: Given guess $\Psi^{(k)}$, solve the dynamic rational inattention problem (DRIP) for each sector $i \in \{1, \ldots, n\}$.
c. **Update price coefficients**: Using the solutions from step 2, compute the new price impulse responses and update $\Psi^{(k+1)}$.
d. **Check convergence**: If $\|\Psi^{(k+1)} - \Psi^{(k)}\| < \epsilon$ for a tolerance $\epsilon > 0$, stop. Otherwise, return to step 2.

B.3. Model's state-space representation

Given a guess $\Psi^{(k)}$, the state space representation for the problem faced by firms in industry i is:

$$x_t = Bx_{t-1} + Qu_t, \tag{A7}$$

$$p_{it}^* = H_i^{(k)} x_t, \tag{A8}$$

where $u_t = (\varepsilon_{mt}, \varepsilon_{z_1,t}, \ldots, \varepsilon_{z_n,t})'$ is the vector of innovations with $u_t \sim \mathcal{N}(0, I_{n+1})$.

The transition matrix $B \in \mathbb{R}^{(n+1)\bar{L} \times (n+1)\bar{L}}$ has block structure:

$$B = \begin{bmatrix} B^m & 0 \\ 0 & B^z \end{bmatrix},$$

where $B^m \in \mathbb{R}^{\bar{L}\times\bar{L}}$ captures the dynamics of the cumulated monetary shock and $B^z \in \mathbb{R}^{n\bar{L}\times n\bar{L}}$ captures the productivity shock dynamics.

The observation equation (A8) links the state to firm i's optimal price:

$$H_i^{(k)} = \alpha_i \Psi_m + \sum_{j=1}^{n} a_{ij} \Psi_j^{(k)} - \Psi_{z_i}, \tag{A9}$$

where Ψ_m denotes the rows of Ψ corresponding to monetary shocks, $\Psi_j^{(k)}$ corresponds to sector j's price, and Ψ_{z_i} corresponds to sector i's productivity shock.

B.4. Solution to the dynamic rational inattention problem

For each sector i, the dynamic rational inattention problem is:

$$\min_{\{a_{it}\}_{t\geq 0}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \left((a_{it} - H_i^{(k)} x_t)^2 + \omega \mathcal{I}(a_{it}; x_t \mid a_{i,t-1})\right)\right], \tag{A10}$$

subject to the state transition (A7) and the marginal cost process (A9).

The solution to this problem yields three objects: i) K_i, the Kalman gain vector, which determines how firms in industry i update beliefs about x_t after receiving a signal; ii) Y_i, the loading of optimal signals on the state x_t; and iii) $\sigma_{i,e}^2$, the variance of the rational inattention noise. Specifically, optimal signals take the form $S_{fit} = Y_i' x_t + e_{fit}, \quad e_{fit} \sim \mathcal{N}(0, \sigma_{i,e}^2)$, and the update of beliefs given a signal realization is determined by K_i.

B.5. Updating the price coefficients

With the solution to the DRIP, we construct the impulse response functions of prices. If p_{it} responds to shock $\varepsilon_{x,t-\tau}$ with coefficient $w_{i,x,\tau}$, the updated guess for Ψ is:

$$\psi_{i,m,\tau}^{(k+1)} = w_{i,m,\tau}, \tag{A11}$$

$$\psi_{i,z_j,\tau}^{(k+1)} = w_{i,z_j,\tau}, \tag{A12}$$

for all sectors i, shock types, and lags $\tau = 0, \ldots, \bar{L} - 1$.

The algorithm then returns to step 2. and iterates until the price coefficients converge.

Appendix C. Additional empirical results

This section complements the empirical sections 2 and 2.3 by providing more details on the data and additional empirical results.

C.1. Survey of Firms' Inflation Expectations (SoFIE)

The SoFIE survey encompasses a diverse range of industries, including nine categories within manufacturing and six within the services sector. Furthermore, it categorizes firms based on their sales-to-GDP ratio and employee count, distinguishing between small (1-19 employees), medium (20-249 employees), and large (250+ employees) enterprises. On average, firms engage in the survey for 3.4 waves, with a standard deviation of 3.5. This structured approach ensures a representative sample that captures the perspectives of businesses across various sectors and scales of operation, providing valuable data on inflation expectations within the corporate landscape.

For the second question on inflation beliefs, there is a rotation across four different questions. Each question is asked once per year. In the first quarter, firms are asked about the probability that inflation one year later will exceed 5%. In the second quarter, they are asked what inflation rate they think the Federal Reserve is targeting on average. In the third quarter, they are asked what they think the inflation rate has been over the last twelve months. In the fourth quarter of each year, they are asked about what they think inflation will be over the next five years on average.

For all questions, respondents are expected to enter a numeric forecast, but the survey instrument accepts ranges (we take the midpoint as the point prediction) or a comment (if possible, we extract a point prediction from the comment). A more detailed description of the survey and its methodology can be found in Federal Reserve of Cleveland (2023).

The relative importance of different sectors and sizes can be adjusted by constructing weights to replicate the distribution of payroll across industries/sizes in the U.S. Candia, Coibion, and Gorodnichenko (2024) argues that this adjustment does not lead to major differences in their results when doing empirical analysis using the SoFIE data.

C.2. Data sources

Table A1 summarizes the variables used in our empirical analysis, their definitions, and data sources.

Figure A1 plots the evolution of industries' forecast errors over the time sample.

TABLE A1. Data Sources and Variable Definitions

Variable	Definition	Source	Period
Inflation expectations	1-year ahead CPI inflation forecast	SoFIE	2018Q2–
Inflation perceptions	Perceived CPI inflation over past 12 months	SoFIE	2018Q3–
Domar weight	Industry sales as fraction of GDP	BEA I-O Tables	2022
Upstreamness	Column sum of Leontief inverse	BEA I-O Tables	2022
Price flexibility	Probability of price readjustment	Pastén, Schoenle, and Weber (2024)	1998–2014
TFP volatility	Std. dev. of linearly detrended log-TFP	BEA/BLS ILPA	1987–2021
CPI inflation	Consumer Price Index, 12-month change	BLS	2018–

TABLE A2. Anchoring Coefficient: Nowcast on Lagged 1-Year Expectation

	$E_t[\pi_t]$
$E_{t-4}[\pi_{t+4}]$	0.242***
	(0.051)
Observations	750
R^2	0.062

Firm-clustered standard errors in parentheses
Time fixed effects included
*** $p < 0.01$

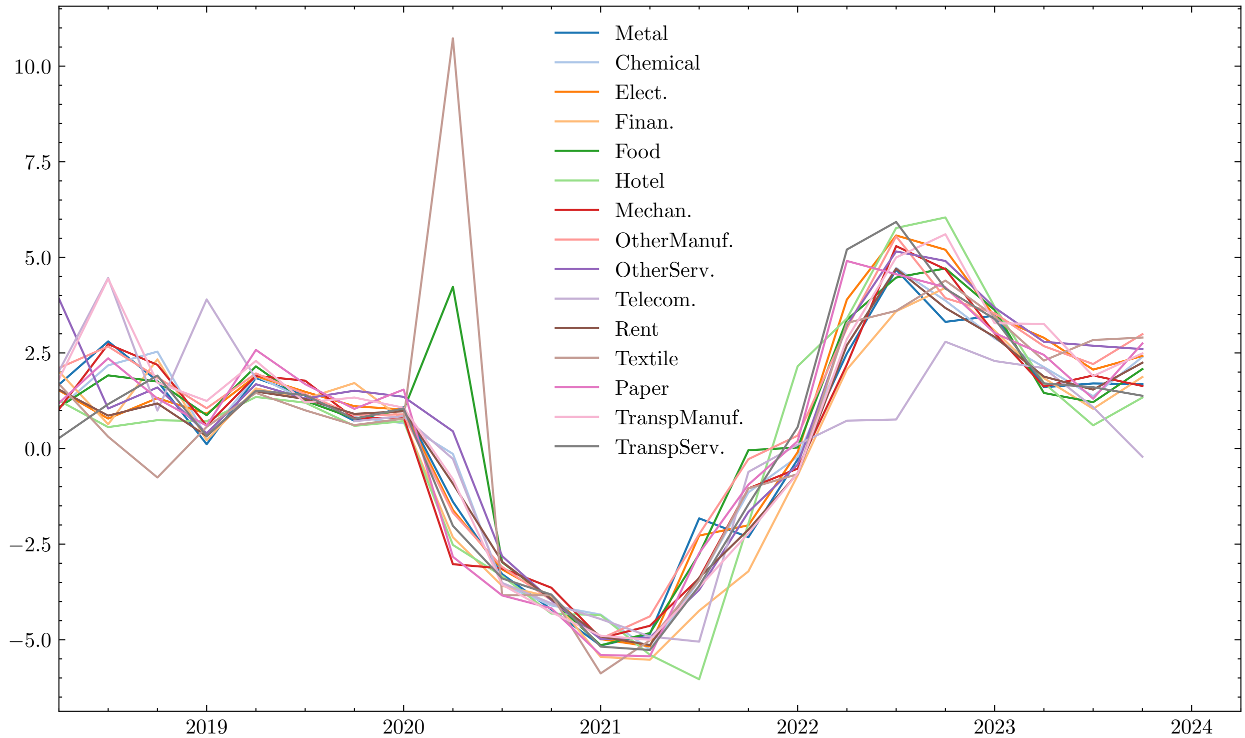
C.3. Calibration of the Information Cost Parameter

Table A2 reports the regression results used to calibrate the information cost parameter ω. Following Afrouzi (2024), we estimate the degree to which firms anchor their inflation forecasts to their prior beliefs. The anchoring coefficient measures how much weight firms place on their prior beliefs when forming current expectations. A higher anchoring coefficient indicates greater reliance on prior beliefs, which corresponds to a higher information cost in the model.

C.4. Robustness: Alternative Domar Weight Specifications

Table A3 presents robustness results for the main empirical specification in Table 1. We include both sales-based and cost-based Domar weights to examine whether different

FIGURE A1. Time-series of one-year-ahead inflation expectations



measures of industry importance affect our conclusions. The results confirm that the sales-based Domar weight remains a significant predictor of forecast accuracy, while the cost-based measure does not add explanatory power. Price flexibility continues to be strongly associated with more accurate inflation forecasts. Upstreamness and TFP volatility remain statistically insignificant across all specifications.

TABLE A3. Forecast Errors and Industry Characteristics (Both Domar Weights)

	$\lvert FE_{0y}\rvert$			$\lvert FE_{1y}\rvert$		
	(1)	(2)	(3)	(4)	(5)	(6)
Domar weight (sales)	-0.729**	-0.639**	-0.639**	-0.287	-0.388**	-0.374**
	(0.331)	(0.314)	(0.315)	(0.205)	(0.162)	(0.163)
Domar weight (cost)	0.158	-0.219	-0.164	0.441*	0.259	0.289
	(0.442)	(0.417)	(0.418)	(0.267)	(0.221)	(0.221)
Upstreamness	0.008	0.014	0.010	-0.028	-0.018	-0.023
	(0.039)	(0.036)	(0.036)	(0.021)	(0.017)	(0.017)
TFP volatility	0.023	0.032	0.030	-0.012	-0.007	-0.009
	(0.033)	(0.031)	(0.031)	(0.020)	(0.016)	(0.016)
Price flexibility	-2.244**	-2.746***	-2.616***	-0.648	-1.397***	-1.227***
	(0.958)	(0.892)	(0.895)	(0.542)	(0.423)	(0.424)
Observations	2,211	2,211	2,211	7,967	7,967	7,967
R^2	0.015	0.137	0.139	0.001	0.353	0.355
Time FE	No	Yes	Yes	No	Yes	Yes
Size controls	Yes	No	Yes	Yes	No	Yes

Note: This table reports OLS regressions of absolute forecast errors on industry characteristics. Includes both sales-based and cost-based Domar weights. $\lvert FE_{0y}\rvert$ corresponds to absolute nowcast error; $\lvert FE_{1y}\rvert$ corresponds to absolute 1-year ahead forecast error. Robust standard errors in parentheses. Size controls correspond to firm size dummies (Medium, Large; reference: Small). ***, **, * indicate statistical significance at 1, 5, and 10 percent, respectively.