Dynamic Pricing under Information Frictions: Evidence from Firm-level Subjective Expectations

Luca Gagliardone* Joris Tielens†

February, 2026

Abstract

This paper develops and implements a structural method to quantify how forward-looking firms are and estimate their effective discount factor from the dynamic pass-through of production-cost shocks into prices. We propose a model featuring real, nominal, and information frictions, and demonstrate how incomplete information distorts the pricing behavior to appear as if firms were myopic, anchoring current prices to past decisions, and underreacting to idiosyncratic cost shocks. Using a unique monthly micro-dataset for the Belgian manufacturing sector that links firm-level prices and costs to survey-based expectations, we find a significant contemporaneous pass-through but a muted response to expected future cost growth. The implied monthly discount factor is about 0.8. These results are consistent with an essentially static pricing behavior, with firms responding strongly as cost shocks materialize. Our micro-estimates imply an "hybrid" Phillips curve with high aggregate discounting and persistence, which can explain the hump-shaped response of inflation to shocks and its long and variable lags.

*Gagliardone: Yale University. Email: luca.gagliardone@yale.edu.

†Tielens: National Bank of Belgium and KU Leuven. Email: joris.tielens@nbb.be.

Luca is greatly indebted to Jess Benhabib, Mark Gertler, Ricardo Lagos, Simone Lenzu, and Venky Venkateswaran for their continued support and advice. We thank Hassan Afrouzi, Michele Andreolli, Isaac Baley, Christoph Boehm, Jaroslav Borovicka, Nicoló Ceneri, Timothy Christensen, Keshav Dogra, Al-Mahdi Esbim, Joel Flynn, Simon Gilchrist, Joao Guerreiro, Zhen Huo, Luigi Iovino, Jennifer La'O, Jean-Paul L'Huillier, Guido Menzio, Tommaso Monacelli, Simon Mongey, Giuseppe Moscarini, Emi Nakamura, Lukas Nord, Mathieu Pedemonte, Diego Perez, Michael Peters, Oliver Pfauti, Pascual Restrepo, Kartik Sastry, Raphael Schoenle, Yatheesan Selvakumar, Antonella Trigari, Aleh Tsyvinski, and Jacob Weber. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the National Bank of Belgium, the Eurosystem, or any other institution with which the authors are affiliated.

1 Introduction

One central yet largely unresolved question in macroeconomics concerns how forward-looking firms are when they set prices, and, in particular, what discount factor they effectively apply to future fundamentals in their pricing decisions. The models most commonly used for quantitative analysis and policy evaluation typically calibrate this parameter close to one, reflecting the prevailing convention that firms behave as rational, highly forward-looking optimizers. Departures from this practice arise in the incomplete-information tradition (Lucas 1972, Woodford 2001, Mankiw and Reis 2002), which maintains rationality but relaxes full information, and in behavioral macroeconomics (e.g., Farhi and Werning 2019, Gabaix 2020), which relaxes strong rationality directly. While these approaches have generated a rich set of theoretical insights, empirical evidence on the discount factor governing firms' price-setting—and, more broadly, on the extent of forward-looking behavior in price dynamics—remains limited and is the subject of active debate.

The scarcity of reliable empirical estimates reflects two related difficulties. First, identifying the discount factor is conceptually and practically very challenging. When using aggregate time-series data, weak instruments, measurement error, and endogeneity concerns often undermine the credibility of standard identification strategies. The growing use of microdata has alleviated some of these issues (McLeay and Tenreyro 2020, Hazell et al. 2022, Gagliardone et al. 2025), but estimating the discount factor remains particularly demanding because it requires information not only on firms' realized pricing decisions but also on their subjective beliefs of future outcomes at the time those decisions are made (Afrouzi et al. 2024). Second, observed pricing behavior features rich dynamics that can reflect multiple forces at once: informational frictions and behavioral distortions, but also nominal and real rigidities that generate inertia in prices and endogenous markups. As a result, a credible empirical strategy must disentangle the forward-looking components from the backward-looking persistence generated by such well-established mechanisms, rather than attributing all inertia or sluggish adjustment to a single channel.

In this paper, we leverage a unique monthly micro-dataset for the Belgian manufacturing sector that combines administrative information on firm-level prices and production costs with survey-based measures of firms' expectations. We use these data to implement a novel structural approach that estimates the discount factor from the dynamic pass-through of cost shocks into prices, explicitly allowing for the role of beliefs. On the theory side, we propose a framework that generalizes the benchmark model with nominal price-setting frictions by incorporating both real rigidities (through endogenous markups) and incomplete information. We fully characterize firms' pricing rules in closed form and derive an explicit mapping from the model to the data in the form of an implementable

moment restriction on the joint intertemporal distribution of prices and costs. This mapping makes it possible to state the identification argument for the firms' discount factor formally and, at the same time, to clarify how information frictions shape the dynamic response of prices to cost shocks. A key implication is an observational equivalence result for the price-setting problem: incomplete information under rationality can generate the same reduced-form firm dynamics as the leading behavioral departures from rationality emphasized in, e.g., Gabaix and Laibson (2017), Bordalo et al. (2020), and Angeletos and Huo (2021). In particular, the pricing behavior of the firm can appear to be myopic—that is, an effective discount factor far below one,— and displaying long-run underreaction to idiosyncratic cost shocks and anchoring of current prices to past decisions.

Our empirical estimates from the dynamic pass-through regression offer a new perspective on how firms set prices. We find that the pass-through of expected future cost shocks is small, which translates into an estimate of the monthly discount factor of approximately 0.8, statistically smaller than the benchmark. At the same time, the pass-through of contemporaneous cost shocks is large, implying—under our maintained assumptions—a slope of the marginal-cost-based Phillips curve of approximately 6%, an order of magnitude larger than influential estimates for the output-based curve (Rotemberg and Woodford 1997, Hazell et al. 2022). Moreover, we find only limited evidence of anchoring and underreaction at the firm level: the estimated coefficients are quantitatively small and, in some specifications, not statistically distinguishable from zero. Taken together, these results are consistent with an environment in which firms set prices almost statically: they respond strongly as cost shocks materialize, while placing relatively little weight on expected future costs and past decisions when setting prices.

Our methodology for solving the model and mapping it to the data not only provides a way to estimate the structural coefficients but also an operational approach to aggregation under heterogeneity. We derive in closed form an "hybrid" Phillips curve (Galí and Gertler 1999), which is able to capture the hump-shaped response of inflation to shocks and its long and variable lags. Our micro-estimates imply, at the aggregate level, a high discounting, which permits our model to generate plausible effects of forward guidance. Finally, we provide an additional validation exercise using aggregate data. We construct an aggregate measure of marginal cost and feed it into the model to obtain the predicted series for inflation, which we compare against the data. Variation in costs can explain more than two-thirds of the volatility of inflation, with a remarkable OLS coefficient of a regression of data on the model of 0.98.

Contribution to the literature. This paper contributes to several strands of the literature.

First, we develop a novel structural approach for estimating firms' dynamic pricing behavior in an environment that jointly features nominal rigidities, real rigidities, and information frictions. We show how to map the primitive assumptions into an empirical specification that nests as special cases the standard cost-price pass-through regressions used in macroeconomics and international macroeconomics (e.g., Auer and Schoenle 2016, Amiti et al. 2019, Amiti et al. 2022, Gagliardone et al. 2025, Ghassibe et al. 2025) as well as in industrial organization (e.g., Goldberg and Hellerstein 2008, Nakamura and Zerom 2010). A key difference from most reduced-form approaches is that we exploit for identification an overidentified restriction on the reduced-form parameters, which is obtained from the mapping between observed pass-through coefficients and structural objects such as the markup elasticity and the persistences of fundamentals and beliefs.

Second, we contribute to the active literature studying the implications of incomplete information for inflation dynamics in models with nominal rigidities. Influential contributions include Nimark (2008), Angeletos and La'O (2009), Maćkowiak and Wiederholt (2009), Hellwig and Venkateswaran (2014), Melosi (2014), Paciello and Wiederholt (2014), Baley and Blanco (2019), and Angeletos and La'O (2020). The closest paper to ours is Angeletos and Huo (2021), which provides an incomplete-information foundation for myopia and anchoring at the aggregate level. Our contribution is complementary in two respects. Theoretically, we derive an observational-equivalence result at the firm level, which leads to a novel empirical specification that takes the form of an error correction model (Engle and Granger 1987) with long-run underreaction. Empirically, we provide a validation of this class of models using firm-level price and cost data combined with direct measures of beliefs, thereby taking the theory's key mechanism—the role of subjective expectations under information constraints—to the data in a novel way.

Finally, we contribute to the growing literature that uses surveys to study firms' expectations and how expectations relate to firms' pricing and inflation dynamics (e.g., Coibion et al. 2020, Andrade et al. 2022, Dogra et al. 2023, Candia et al. 2024). A distinctive feature of our setting is that survey-based expectations can be matched directly to firms' realized pricing decisions and production costs, allowing us to assess how expectations affect pricing decisions. In line with Coibion et al. (2018), we find that firms' reported expectations exhibit patterns consistent with rational behavior: firms tend to correctly anticipate whether they are more likely to raise or lower prices in the near future.

2 Theoretical framework

This section presents the theoretical framework that underlies our empirical analysis. We formulate the minimal structure required to produce firm pricing equations and identify the structural parameters of the economy, which pin down the firms' effective discount factor and the aggregate dynamics of inflation.

2.1 Primitives

The economy is populated by a continuum of heterogeneous firms indexed by $i \in [0,1]$. Firms compete monopolistically within an industry, and take industry and economy-level prices as given. In each period t, firms set their log price $p_{i,t} \equiv \ln P_{i,t}$ subject to a downward-sloping demand curve á la Kimball (1995). Prices are set dynamically to maximize expected intertemporal profits subject to a nominal rigidity á la Calvo (1983), where expectations are formed on the basis of possibly imperfect information regarding the realization and evolution of the firm's real marginal cost.

Household's preferences. We assume that aggregate consumption Y_t is a bundle of differentiated intermediate inputs $Y_{i,t}$. The bundle of inputs is assembled into a final good using the Kimball aggregator:

$$\int_0^1 \Upsilon\left(\frac{Y_{i,t}}{Y_t}\right) di = 1 \tag{1}$$

where $\Upsilon(\cdot)$ is a continuous and twice-differentiable function, strictly increasing and strictly concave, and homogeneous of degree one with $\Upsilon(1) = 1$. This flexible aggregator provides a tractable benchmark that captures—for our purposes—the key quantitative features of several homothetic demand systems that depart from a constant elasticity of substitution. Under CES, markups are constant; more generally, Kimball allows the elasticity to vary with the sales share, providing a tractable foundation for incomplete long-run pass-through of idiosyncratic cost shocks and time-varying markups.

Taking the aggregate demand Y_t as given, the problem of the household is to minimize expenditure for each differentiated good subject to the aggregate constraint:

$$\min_{Y_{i,t}} \int_0^1 P_{i,t} Y_{i,t} di \qquad \text{s.t.} \quad \int_0^1 \Upsilon\left(\frac{Y_{i,t}}{Y_t}\right) di = 1.$$

The solution of the household problem leads to a downward-sloping demand curve for the good $Y_{i,t}$, which is of the form:

$$Y_{i,t} = d(p_{i,t}, p_t) Y_t, \tag{2}$$

where $d(p_{i,t}, p_t) \equiv (\Upsilon')^{-1}\Big(\Upsilon'(1)\exp(p_{i,t} - p_t)\Big)$ is a strictly decreasing function of relative prices, and $p_t \equiv \ln P_t$ is the log aggregate price index consistent with equation (1).[1]

Fundamental. Denote by $TC_{i,t} \equiv TC(Y_{i,t})$ the total nominal cost of producing $Y_{i,t}$ units of output and by $MC^n_{i,t} \equiv \partial TC_{i,t}/\partial Y_{i,t}$ the nominal marginal cost of production. Let $mc^n_{i,t} \equiv \ln MC^n_{i,t}$ be the log nominal marginal cost and $mc_{i,t} \equiv mc^n_{i,t} - p_t$ be the log real marginal cost. We consider a production function with constant returns to scale (relaxed in the empirical analysis) and suppose that the real marginal cost follows an exogenous autoregressive process of order one with persistence parameter ρ:[2]

$$mc_{i,t} = \rho mc_{i,t-1} + \varepsilon_{i,t}, \quad \varepsilon^{\xi}_{i,t} \equiv (\varepsilon_{i,t} - \varepsilon^a_t) \sim \mathcal{N}(0,\ \sigma^2_\xi), \quad \varepsilon^a_t \sim \mathcal{N}(0,\ \sigma^2_a), \quad |\rho| < 1 \tag{3}$$

The real marginal cost is subject to two forms of fundamental disturbances. The error term $\varepsilon^{\xi}_{i,t}$ is a firm-specific shock distributed in the cross-section according to a Gaussian distribution with mean zero and variance $\sigma^2_\xi > 0$. The shock ε^a_t is a Gaussian aggregate shock with mean zero and variance $\sigma^2_a > 0$. Thanks to the continuum of firms, we can assume that the two disturbances are uncorrelated, implying $\varepsilon^{\xi}_{i,t} \perp \varepsilon^a_t$ by normality. We interpret the idiosyncratic disturbance as a firm-specific real marginal-cost shifter, such as total-factor productivity. The aggregate disturbance captures correlated real shocks to firms' costs, such as an aggregate TFP shock. We denote by $\sigma^2_\varepsilon \equiv \sigma^2_a + \sigma^2_\xi$ the variance of $\varepsilon_{i,t}$. We provide empirical evidence consistent with an AR(1) and discuss robustness to steady-state heterogeneity and decreasing returns to scale in section 4.3.

Information. We suppose that firms face a systematic measurement error of their real marginal cost. This information friction may arise, for example, because firms cannot estimate precisely their production function (thus observe $mc^n_{i,t}$ with noise), or because they have imperfect knowledge of the aggregate price level. We model the measurement error using a private, noisy signal $s_{i,t}$. Without loss of generality, we normalize the signal to correspond to the conditional mean of the real marginal cost:

$$s_{i,t} = mc_{i,t} + u_{i,t}, \qquad u_{i,t} \sim \mathcal{N}(0, \sigma^2_u), \qquad u_{i,t} \perp \varepsilon_{i,t},\ \varepsilon^a_t \tag{4}$$

[1] The demand system depends on relative prices, defined as firm-level minus the industry price index (competitors' prices). Real variables are defined in deviation from the economy-level price level. We simplify the notation by having only one industry in the theoretical section, as it is immaterial for our purposes: the empirical analysis includes industry-by-time fixed effects, hence controls for both industry and economy-level aggregates.

[2] We normalize the steady-state real marginal cost to be equal to zero. This normalization is immaterial for our empirical analysis (it only affects the interpretation of the intercept), but simplifies the notation. We relax this assumption in section 4.3.

As a limit case, when the private signal is perfectly informative ($\sigma_u^2 \to 0$), firms observe their real marginal cost without noise. The orthogonality restriction between measurement error and fundamental innovations, together with linearity of the process and Gaussian distribution of shocks, allows firms to optimally forecast the dynamics using the Kalman filter. We illustrate these points in section 2.2. Finally, we require that prior beliefs at time zero are normally distributed.

We denote by $\mathcal{I}_{i,t} \equiv \{s_{i,\tau}\}_{\tau \leq t}$ the information set of firm i at time t and denote by $\mathbb{E}_{i,t}(\cdot) \equiv \mathbb{E}(\cdot|\mathcal{I}_{i,t})$ the subjective expectation conditional on the information set. Our empirical analysis can allow for observability of other variables as well, with some caveats. We discuss extensions and robustness to alternative information structures in section 4.3.

Law of motion of prices. Price changes are allowed to be infrequent and governed by a Bernoulli process á la Calvo. Denote by $p_{i,t}^o \equiv \ln P_{i,t}^o$ the log "reset" prices, i.e., the price that a firm would like to set conditional on adjusting. The reset price is chosen by firms to maximize the expected present value of profits, as discussed in detail in the next section. The law of motion of realized price changes ($\Delta p_{i,t} \equiv p_{i,t} - p_{i,t-1}$) is given by:

$$\Delta p_{i,t} = \mathbb{I}_{i,t} \cdot (p_{i,t}^o - p_{i,t-1}) \tag{5}$$

where $\mathbb{I}_{i,t} \overset{i.i.d.}{\sim} \mathcal{B}(1-\theta)$ is a Bernoulli random variable with constant intensity $(1-\theta) \in [0,1)$, which denotes the frequency of price changes.

2.2 Firm's problem

The problem of the firm is to maximize the expected present value of current and future real profits, subject to the demand curve in (2) and the information and nominal frictions. Formally, firms choose a constant log reset price $p_{i,t}^o \equiv \ln P_{i,t}^o$ (conditional on adjusting $\mathbb{I}_{i,t} = 1$) to maximize the following objective function for every $t \geq 0$:

$$\begin{aligned} \max_{P_{i,t}^o} \ & \mathbb{E}_{i,t} \left\{ \sum_{\tau \geq 0} \theta^\tau \left[\Lambda_{t,\tau} \left(\frac{P_{i,t}^o}{P_{t+\tau}} \ Y_{i,t+\tau} - \frac{TC(Y_{i,t+\tau})}{P_{t+\tau}} \right) \right] \right\} \\ \text{s.t.} \quad & Y_{i,t+\tau} = d(p_{i,t}^o, p_{t+\tau}) Y_{t+\tau} \end{aligned} \tag{6}$$

where $\Lambda_{t,\tau}$ denotes the stochastic discount factor between t and $t+\tau$. Differentiating the objective function with respect to $P_{i,t}^o$ leads to the following first-order condition:

$$\mathbb{E}_{i,t} \left\{ \sum_{\tau \geq 0} \theta^\tau \Lambda_{t,\tau} Y_{i,t+\tau} \left[\frac{P_{i,t}^o}{P_{t+\tau}} - (1 + \mu_{i,t+\tau}) \frac{MC_{i,t+\tau}^n}{P_{t+\tau}} \right] \right\} = 0 \tag{7}$$

where $\mu_{i,t}$ denotes the desired markup, which corresponds to the Lerner index and is a function of the residual elasticity of demand $\epsilon_{i,t}$:

$$\mu_{i,t} = \ln\left(\frac{\epsilon_{i,t}}{\epsilon_{i,t}-1}\right), \qquad \epsilon_{i,t} \equiv -\frac{\partial \ln Y_{i,t}}{\partial p^o_{i,t}} \tag{8}$$

The reset price of a firm is the expected discounted present value of the current and future markups over marginal costs. As anticipated, the demand curve á la Kimball implies that the residual elasticity of demand $\epsilon_{i,t}$ can be time-varying and firm-specific. When this is the case, the markup is time-varying and firm-specific. As shown in Appendix A.1, the elasticity depends on the firm sales shares ($r_{i,t} \equiv Y_{i,t}/Y_t$) through the Kimball aggregator via $\epsilon_{i,t} = -\Upsilon'(r_{i,t})/(\Upsilon''(r_{i,t}) \cdot r_{i,t})$. This implies that firms with large sales shares have higher markups, a salient feature of the data.

Log-linearization. In order to obtain a linear representation of the optimality conditions of the model that can be mapped to the data, we log-linearize the first-order condition (7) and the markup (8) around the symmetric steady state.[3] The log-linearized markup depends inversely on relative prices and is given by:

$$\mu_{i,t} = -\Gamma(p^o_{i,t} - p_t), \quad \Gamma \geq 0. \tag{9}$$

where $\Gamma \equiv -\frac{\epsilon'}{\epsilon(\epsilon-1)}\frac{\Upsilon'(1)}{\Upsilon''(1)}$ is the elasticity of the markup to relative prices, which depends on the derivative of the residual elasticity of demand at the steady state (ϵ') as well as on the steady-state elasticity (ϵ). Log-linearizing the first-order condition and replacing the log-linearized markup, we obtain the following recursive formulation for the reset price:

$$p^o_{i,t} = \mathbb{E}_{i,t}\Big\{(1-\beta\theta)\Big[(1-\Omega)mc^n_{i,t} + \Omega p_t\Big] + \beta\theta p^o_{i,t+1}\Big\} \tag{10}$$

where $\Omega \equiv \frac{\Gamma}{1+\Gamma} \in [0,1)$ is the degree of strategic complementarities and β is the household's discount factor ($\Lambda_{t,1}$ at the steady state). Equation (10) is the familiar log-linearized pricing equation under Calvo pricing, extended to allow for strategic complementarities ($\Omega \neq 0$) and incomplete information. In log-deviation from the steady state, the log reset price can be written as the expected discounted present value of a linear combination of the firm's marginal cost ($mc^n_{i,t}$) and the aggregate price index (p_t). The discounting of the expectation is given by $\beta\theta$, as it takes into account the probability that the price remains fixed between any two consecutive periods (θ).

[3] The symmetric steady state (relative prices equal to one) ensures that the elasticity of the markup at the steady state is constant across firms, so that the coefficient of the expansion (Γ) is not firm-specific. We discuss relaxations of this assumption in section 4.3.

Micro vs macro dynamics. Equation (10) shows that, in order to reset prices, firms need to compute the expected marginal cost $\mathbb{E}_{i,t}(mc^n_{i,t+h})$, $h \geq 0$. We now show that the problem of the firm can be neatly separated into the forecast of a real common component and an idiosyncratic component, defined respectively as:

$$mc_t \equiv \int_0^1 mc_{i,t}\, di, \qquad \xi_{i,t} \equiv mc_{i,t} - mc_t$$

To show this, we take expectations of the law of motion for a price change (equation 5) and combine it with the log-linearized first-order condition (derivations are in Appendix A.2). Denoting by $\pi_t \equiv p_t - p_{t-1}$ the aggregate inflation rate and by $\widehat{p}_{i,t} \equiv p_{i,t} - p_t$ relative prices, we decompose additively expected price changes in $\mathbb{E}_{i,t}(\Delta p_{i,t}) = \mathbb{E}_{i,t}(\pi_t) + \mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t})$. The two components satisfy the respective pricing equations:

$$\mathbb{E}_{i,t}(\pi_t) \;=\; \mathbb{E}_{i,t}\Big\{\kappa(1-\Omega)mc_t + \beta\pi_{t+1}\Big\}, \qquad \kappa \equiv \frac{(1-\beta\theta)(1-\theta)}{\theta} \tag{11}$$

$$\mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t}) \;=\; \mathbb{E}_{i,t}\left\{\kappa\Big[(1-\Omega)\xi_{i,t} - \widehat{p}_{i,t}\Big] + \beta\Delta\widehat{p}_{i,t+1}\right\} \tag{12}$$

This decomposition shows that, despite sharing the same structural parameters, aggregate inflation and relative prices do not obey the same expected dynamics. Whereas equation (11) is reminiscent of a standard Phillips curve relationship with subjective expectations, relative prices in equation (12) satisfy an error-correction model (Engle and Granger 1987), according to which short-term fluctuations in costs lead to persistent deviations in prices. The term in square brackets ($(1-\Omega)\xi_{i,t} - \widehat{p}_{i,t}$) captures deviations from the long-run cointegrating relationship between prices and costs, which is obtained by setting the term equal to zero. Whenever an idiosyncratic cost shock occurs ($\xi_{i,t}$), in the long run, a fraction $(1-\Omega)$ is passed into the firm price, maintaining the cointegrating relationship. However, in the short run, the adjustment to the relative price level ($\widehat{p}_{i,t}$) is frictional, generating a temporary departure from the cointegrating relationship. For a positive cost shock ($\xi_{i,t} \geq 0$), the relative price adjusts in the short run less than $(1-\Omega)\xi_{i,t}$, so that the term in square brackets becomes positive. Since $\kappa > 0$, this leads, in expectations, to upward pressure on the growth rate of relative prices. As prices are forward-looking, part of the adjustment in relative prices is anticipated, speeding up the convergence.

A central feature of the model is that, as aggregate inflation and relative prices follow different dynamics, information remains incomplete even if firms observed directly their real marginal cost ($\sigma^2_\varepsilon/\sigma^2_u \to \infty$). Indeed, they would still not be able to dynamically disentangle the idiosyncratic and the aggregate components of real costs. Therefore, a

more informative signal alleviates but does not generally resolve the information friction. As we will see in section 5, it is precisely this feature that leads to a rich—and realistic—hump-shaped dynamics of aggregate inflation.

In the remaining part of the section, we characterize the behavior of relative prices. This is aimed at drawing the mapping between the theoretical model and the empirical specification, which includes time fixed effects and therefore estimates the structural parameters of the model from cross-sectional variation (equation 12). We complete the characterization of equation (11) and clarify how the micro and macro dynamics relate when deriving the aggregate implications of the model in section 5.

Kalman filter. The next step is to spell out how firms form expectations of costs using the information available. Under the assumptions of linear dynamics with Gaussian shocks (and prior beliefs), the optimal update of beliefs is computed recursively via the Kalman filter. Starting from the initial belief $\mathbb{E}_0(mc_{i,0})$, beliefs are updated (in a closed loop) upon receiving the new information $s_{i,t}$ as:

$$\mathbb{E}_{i,t}(mc_{i,t}) \;=\; \lambda\, \mathbb{E}_{i,t-1}(mc_{i,t-1}) + \left(1 - \frac{\lambda}{\rho}\right) s_{i,t}, \tag{13}$$

where the persistence of beliefs corresponds to the stabilizing root of the Riccati equation, which satisfies $0 < \lambda < \rho$ and is given by (derivations in Appendix A.3):

$$\lambda = \frac{1}{2}\left[\rho + \frac{1}{\rho}\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_u^2}\right) - \sqrt{\left(\rho + \frac{1}{\rho}\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_u^2}\right)\right)^2 - 4}\right]$$

The Kalman gain is the weight on the new information, given by $1 - \lambda/\rho$. When the signal-to-noise ratio converges to infinity ($\sigma_\varepsilon^2/\sigma_u^2 \to \infty$), the Kalman gain converges to one ($\lambda \to 0$) as the signal fully reveals the shock.

Because the two components ($\xi_{i,t}$ and mc_t) have the same persistence (ρ), firms cannot disentangle them dynamically.[4] The forecast of each component is then obtained by rescaling by the ratio of variances (σ_ξ^2 and σ_a^2). Defining the rescaled private signal by $\widehat{s_{i,t}} \equiv (\sigma_\xi^2/\sigma_\varepsilon^2) s_{i,t}$, the belief of the idiosyncratic component evolves as:

$$\mathbb{E}_{i,t}(\xi_{i,t}) \;=\; \lambda\, \mathbb{E}_{i,t-1}(\xi_{i,t-1}) + \left(1 - \frac{\lambda}{\rho}\right) \widehat{s_{i,t}}. \tag{14}$$

[4] This is immediate to show from equation (3). Averaging in the cross-section leads to $mc_t = \rho mc_{t-1} + \varepsilon_t$. Subtracting the cross-sectional mean leads to $\xi_{i,t} = \rho \xi_{i,t-1} + \varepsilon_{i,t}^{\xi}$. See Table 4 for empirical estimates of the persistences of the cost components.

Away from the complete-information benchmark, beliefs are slowly updated in light of the new information, leading to additional stickiness in prices on top of that attributable to the nominal rigidity. We clarify this point further in the next section.

2.3 The dynamics of relative prices

The characterization of the relative price-setting problem is obtained by combining equation (12) with the update of beliefs in equation (14). The next proposition illustrates how, under incomplete information and nominal rigidities, the expected price growth displays a rich dynamics that follows an ARMA(2,1) process in the private signals, i.e., a process with a moving average component and two autoregressive lags, which depend respectively on the strength of the two frictions.

Proposition 1 (Dynamics of relative prices)**.** *The growth rate of relative prices satisfies:*

$$\mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t}) = \underbrace{\kappa(1-\Omega)\left(\frac{1}{\vartheta-\rho\beta}\right)\left(1-\frac{\lambda}{\rho}\right)}_{\textit{Constant}} \underbrace{\frac{(1-\mathsf{L})}{(1-\mathsf{L}/\vartheta)(1-\lambda\mathsf{L})}}_{\textit{ARMA(2,1)}} \widehat{s}_{i,t} \tag{15}$$

where L *denotes the lag operator and the root of the AR component is:*

$$\vartheta \equiv \frac{(1+\beta+\kappa)+\sqrt{(1+\beta+\kappa)^2-4\beta}}{2} > 1$$

The proof of the proposition is obtained by guessing that the expected growth rate of relative prices can be written as a causal lag polynomial of the history of the signals and solving the fixed point in equation (12) to verify the guess. These computations require applying the Wiener-Hopf prediction formula to annihilate the non-causal component of expected prices when taking expectations at time t. The verification step involves appropriately choosing the boundary conditions to ensure that the solution is causal (analytical). We refer to Appendix A.4 for the derivations.

Proposition 1 shows that the process for the expected growth rate of relative prices satisfies as a causal representation in terms of the history of private signals with finitely many lags. The moving average component of the ARMA(2,1) process has a unit root, which comes from differencing prices. The autoregressive part has poles given by ϑ, which depends on the strength of the nominal friction via κ, and $1/\lambda$, which is the inverse of the root of the belief process in (14). The multiplicative constant depends on the degree of strategic complementarities: the weaker the strategic complementarities (small Ω), the higher the expected growth of relative prices. Moreover, both the Kalman gain $(1-\lambda/\rho)$

and the slope of the pricing equation (κ) also enters the expression. Therefore, in this model, whereas the real rigidity only affects the volatility of expected price changes via the multiplicative constant, nominal and information frictions affect both the volatility and the persistence (i.e., the roots of the ARMA process).

To further clarify the effects of nominal and information frictions, we take the limit for the signal-to-noise ratio going to infinity and obtain the expected dynamics of prices when the signal is perfectly informative:

$$\lim_{\sigma_\varepsilon^2/\sigma_u^2 \to \infty} \mathbb{E}_{i,t}(\Delta \widehat{p}_{i,t}) = \kappa(1-\Omega)\left(\frac{1}{\vartheta - \rho\beta}\right)\left(\frac{1-\mathsf{L}}{1-\mathsf{L}/\vartheta}\right)\widehat{s}_{i,t}$$

In this case, the volatility of expected price changes is always larger than the incomplete information case for any given realization of the signals (the constant is larger), and the second lag of the autoregressive component becomes irrelevant (ARMA(1,1)), so that prices display less persistence. Complementarily, taking the limit for vanishing nominal rigidity, we obtain:

$$\lim_{\theta \to 0} \mathbb{E}_{i,t}(\Delta \widehat{p}_{i,t}) = (1-\Omega)\left(1 - \frac{\lambda}{\rho}\right)\left(\frac{1-\mathsf{L}}{1-\lambda\mathsf{L}}\right)\widehat{s}_{i,t}$$

Intuitively, absent the nominal rigidity, prices are less persistent (ARMA(1,1)). However, the multiplicative constant is now larger as long as $\kappa < (\vartheta - \rho\beta)/\vartheta$, which is the empirically relevant case, implying that the volatility increases. This comparison shows that, in this model, incomplete information has a similar effect to the nominal friction on the persistence of prices, but has the opposite effect on the volatility. This intuition is key for understanding why it is possible to identify the two forces from data: fixing Ω, for any given level of empirical price persistence and volatility, there is only one combination of the nominal and information frictions that is able to match both. However, the coefficient Ω also scales the volatility of relative price changes.

In the next sections, we show how to overcome these difficulties in estimating the structural parameters of the economy under incomplete information. The identification argument leverages an overidentifying restriction that permits disentangling the persistence in prices due to the backward-looking and the forward-looking components.

2.4 The observationally equivalent representation

In the previous section, we provided a characterization of the structural relation between prices and information upon which firms base their pricing decisions. However, in the data, we do not observe the firm's private signal. In this section, we show how to derive an implementable reduced-form representation of the firm's pricing behavior.

The central idea for the reduced-form representation is to project prices onto the

history of realized costs, and impose that the resulting process preserves all the dynamic properties of the structural representation from proposition 1. Accordingly, without explicitly observing the realized signals, the econometrician would not be able to dynamically distinguish between the structural and reduced-form representations—i.e., they are observationally equivalent. Additionally, this mapping offers an intuitive interpretation: we recast the original economy under incomplete information into an alternative "behavioral" economy in which information is complete, but preferences are distorted.[5] The distortions ensure that the two economies display the same dynamics, hence embed the effects of the subjective uncertainty. This provides not only a way of recovering the structural parameters from data and interpreting the role of incomplete information, but also—as we shall see in section 5— an operational approach to aggregation under heterogeneity.

Formally, denote by $h^{\text{CI}}(\mathsf{L})$ a reduced-form causal lag polynomial that maps realized costs into expected price changes. By observational equivalence, we mean that $h^{\text{CI}}(\mathsf{L})$ is the projection onto the history of costs:

$$\mathbb{E}\Big(\mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t}) \,\Big|\, \{\xi_{i,\tau}\}_{\tau\leq t}\Big) \;=\; h^{\text{CI}}(\mathsf{L})\xi_{i,t} \tag{16}$$

The next proposition derives an observationally equivalent representation.[6]

Proposition 2 (Observational equivalence)**.** *There exist reduced-form parameters $\omega_p, \omega_f, \omega_\ell$ such that relative prices satisfy the following dynamics:*

$$\mathbb{E}_t(\Delta\widehat{p}_{i,t}) = \mathbb{E}_t\left\{\kappa\Big[(1-\Omega)\xi_{i,t} - \omega_p\,\widehat{p}_{i,t}\Big] + \omega_f\Delta\widehat{p}_{i,t+1} + \omega_\ell\Delta\widehat{p}_{i,t-1}\right\} \tag{17}$$

where $\mathbb{E}_t(\cdot)$ denotes the expectation under the complete information set. The coefficients that satisfy the observational equivalence condition (16) *are uniquely given by:*

$$\begin{aligned}
\omega_p &= \frac{(1-\lambda)(1-\vartheta)}{\kappa\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)D}\left\{\lambda\rho(\vartheta-1)+\lambda+\rho^2-\beta\rho^2\Big(\lambda+\frac{1}{\vartheta}\Big)+\Big(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}-1\Big)(\lambda+\rho^2-\rho-\lambda\rho)\right\},\\
\omega_f &= \frac{1}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)D}\left\{\lambda\vartheta\Big[\rho(\vartheta-1)+1\Big]+\beta\rho^2(\lambda-1-\lambda\vartheta)+\Big(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}-1\Big)\vartheta(\lambda-\rho)\right\},\\
\omega_\ell &= \frac{\lambda\rho^2(\beta\lambda-\vartheta)+\left(1-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\lambda\rho\,\vartheta(\rho-\lambda)}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\vartheta\,D}, \qquad D \;\equiv\; (\rho-\lambda)(\lambda\rho-\lambda-\rho-\lambda\rho\vartheta)
\end{aligned} \tag{18}$$

[5] As standard, complete information means that the distribution of fundamentals (and, therefore, aggregate prices) up to time t is observed and common knowledge. We denote by $\mathcal{I}_t \equiv \cup_i \mathcal{I}_{i,t}$ the complete information set and by $\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot|\mathcal{I}_t)$ the expectation. Then, $\mathbb{E}_{i,t}(mc_{i,t+h}) = \mathbb{E}_{j,t}(mc_{i,t+h}) = \mathbb{E}_t(mc_{i,t+h})$, $j, i \in [0,1]$, $h \geq 0$.

[6] The projection is generally not unique. We pick a representation that is minimal in the sense that it has the smallest number of wedges (distorted coefficients) needed to satisfy the restriction.

The proof of the proposition is in Appendix A.5. Comparing it with the pricing equation of the economy under incomplete information (12), three differences stand out. First, the long-run cointegrating relation between prices and cost is distorted by the coefficient ω_p; second, the effective discounting of future price changes is distorted by ω_f; third, there is an additional lagged price change governed by ω_ℓ. The additional lag term and the reweighted forward-looking and long-run components summarize the behavioral distortions needed to replicate the effects of incomplete information. The representation (17) need not have—by itself— a structural interpretation. Nevertheless, it can be estimated with data and used to recover the structural parameters of the model by inverting the mapping from the reduced-form coefficients. Indeed, these are pinned down by three structural parameters: the price stickiness (θ, which governs the slope κ and ϑ), the persistence of beliefs (λ), and the persistence of the cost shocks (ρ).

The next corollary further leverages the reduced-form representation to illustrate the effects of incomplete information and link it to the terminology used in the literature.

Corollary 1 (Pricing behavior)**.** *Under incomplete information ($\sigma_\varepsilon^2/\sigma_u^2 < \infty$), the firms' pricing decisions appear as if they are subject to the following three behavioral distortions:*

- ***Myopia***: $\omega_f < \lim_{\sigma_\varepsilon^2/\sigma_u^2 \to \infty} \omega_f \leq \beta.$
- ***Underreaction***: $\omega_p > \lim_{\sigma_\varepsilon^2/\sigma_u^2 \to \infty} \omega_p \geq 1.$
- ***Anchoring***: $\omega_\ell > \lim_{\sigma_\varepsilon^2/\sigma_u^2 \to \infty} \omega_\ell = 0.$

The proof is in Appendix A.6. When firms are rational and process available information optimally, incomplete information distorts observed behavior in a way that is isomorphic to a complete-information model with behavioral wedges. The proposition connects such distortions to the standard terminology used in the literature. First, because ω_f is generally smaller than the household's discounting β, firms place less weight on expected future price changes than they would under complete information; this appears as myopic behavior (Angeletos and Lian 2018, Angeletos and Huo 2021).[7] Second, the long-run pass-through of idiosyncratic cost shocks is $(1-\Omega)/\omega_p$, which is smaller than the complete-information benchmark $1-\Omega$; we refer to this as underreaction (Bordalo et al. 2020). Third, incomplete information introduces additional persistence in the price dynamic, effectively anchoring current prices to past price changes ($\omega_\ell > 0$). Notably, while

[7] Even when the signal is perfectly informative ($\sigma_\varepsilon^2/\sigma_u^2 \to \infty$), information is not complete as firms cannot perfectly disentangle between $\xi_{i,t}$ and mc_t. In particular, $\omega_f \to \beta$ if and only if information is complete, i.e., when both $\sigma_\varepsilon^2/\sigma_u^2 \to \infty$ and $\sigma_\xi^2/\sigma_\varepsilon^2 \to 1$. The same logic applies to ω_p. See the Appendix A.6 for the details.

myopia and anchoring generate dynamic distortions that fade as shocks dissipate, underreaction implies a permanently dampened response of prices to idiosyncratic cost shocks, and thus plays a role analogous to stronger strategic complementarities.

2.5 Identification

We now describe how to identify the structural parameters of the model.[8] The identification strategy implements equation (17) using the generalized method of moments with one overidentifying parameter restriction. We emphasize that the orthogonality of the instruments follows directly from the structural representation and imposing the overidentifying restriction on the reduced form. The next corollary formalizes the mapping of the model to the data.

Corollary 2 (Orthogonality condition)**.** *Let $\boldsymbol{\Theta} \equiv [\theta,\ \Omega,\ \lambda,\ \rho]$ collect the structural parameters to be estimated and let $Z_{i,t} \in \mathcal{I}_t$ be any random variable measurable with respect to the full information set at time t. The following orthogonality condition holds:*

$$\mathbb{E}\left\{ Z_{i,t} \cdot \widehat{v}_{i,t+1}(\boldsymbol{\Theta}) \right\} = 0, \tag{19}$$

where the residual is:

$$v_{i,t+1}(\boldsymbol{\Theta}) \equiv \Delta p_{i,t} - \kappa \Big[(1-\Omega) mc^{n}_{i,t} - \omega_p\, p_{i,t} \Big] - \omega_f \Delta p_{i,t+1} - \omega_\ell \Delta p_{i,t-1}.$$

$\widehat{v}_{i,t+1} \equiv v_{i,t+1} - \int_0^1 v_{i,t+1}\, di$ denotes the deviation of the residual from its cross-sectional mean, and the reduced-form coefficients $\omega_p, \omega_f, \omega_\ell$ jointly satisfy one overidentifying restriction (defined by the system of equations 18).

The corollary leverages the observationally equivalent representation (proposition 2), which imposes a moment restriction of the joint distribution of prices and costs, conditional on the full information set. By the law of iterated expectations, any variable in $\mathcal{I}_t$ is a suitable candidate for the instrument set as it satisfies orthogonality, provided that the reduced-form coefficients are constrained by the overidentifying restriction.

Let us now heuristically discuss which instruments provide variation that identifies the structural parameters. The frequency of price adjustments, θ, is identified through the slope coefficient, κ, provided that the shifters induce transitory departures from the long-run (cointegrating) relationship between relative prices and the relevant cost component,

[8] We treat the household's discount factor β and the variance ratio $\sigma^2_\xi/\sigma^2_\varepsilon$ as known. The discounting has a standard calibration of approximately one. The variance ratio is pinned down from the data, given an estimated value for the persistence ρ. We discuss further the calibration in section 4.

i.e., fluctuations in $(1-\Omega)mc^n_{i,t} - \omega_p\, p_{i,t}$. Conditional on κ, the degree of strategic complementarities, Ω, is then identified from movements in marginal cost, because Ω governs how strongly firm-level price changes respond to contemporaneous cost variation relative to the endogenous price component. Turning to dynamics, variation in lagged prices identifies the anchoring parameter ω_ℓ. Finally, variation in expected future price changes that is not spanned by lagged prices identifies the effective discount factor, ω_f. These two coefficients jointly pin down the structural parameters (λ, ρ), which govern the persistence of the beliefs and the persistence of costs. The remaining parameter, ω_p, is then pinned down by the model's overidentifying restriction implied by the nonlinear mapping from $\boldsymbol{\Theta}$ to $(\Omega, \kappa, \omega_p, \omega_f, \omega_\ell)$. This discussion motivates an instrument set that combines (i) lagged marginal cost (to shift contemporaneous costs), (ii) lagged firm prices (to shift lagged price changes and relative-price dynamics), and (iii) subjective expectations (to shift anticipated future price changes using information available to firms at time t). The first two components deliver variation that is sufficient to identify κ, Ω, and ω_ℓ, while expectations are essential for identifying ω_f by isolating movements in future prices that are anticipated by firms but not mechanically forecastable from lagged realized prices. We present the set of instruments in section 4.1 after discussing data and measurement.

3 Data and measurement

This section assembles a micro-level dataset that covers the manufacturing sector in Belgium between 1995 and 2023 at a monthly frequency. We begin by outlining the various data sources and subsequently discuss the mapping of the theoretical counterparts from section 2 to the data. Finally, we discuss various features of the data that justify aspects of the theoretical framework and are salient for the identification of the model.

3.1 Data sources

Our analysis draws on four confidential data sources and combines detailed firm-level information on monthly production, costs, pricing decisions, and subjective expectations on future firm-level variables. Our dataset extends and enriches the quarterly dataset used by Gagliardone et al. (2025, 2024) to include data on firm beliefs.

First, the PRODCOM survey tracks domestic firms' monthly sales and physical quantities sold for each narrowly defined (8-digit product codes) manufacturing product (e.g., Pineapple juice, vs Grape juice, vs Apple juice). The PRODCOM sampling strategy is designed to cover at least 90% of domestic production value within each manufacturing industry by surveying all firms operating in the country with (i) a minimum of 20 employees

or (ii) total revenue above 4.5 million euros. We use this information to calculate domestic unit values (sales over quantities) at the firm-product level.[9] Second, we use detailed administrative data to measure firms' variable production costs. Specifically, we obtain information on firms' monthly purchases of intermediates (materials and services) from their value-added tax (VAT) declarations submitted to the fiscal authorities. Additionally, we draw upon firms' social security declarations to obtain a measure of their labor costs (the wage bill). Finally, we tap into the monthly Business Survey of the National Bank of Belgium (NBB-BS). This survey tracks a representative panel of the Belgian manufacturing sector and is designed to provide both a timely and bottom-up measure of various business cycle indicators (De Greef and Van Nieuwenhuyze 2009). The NBB-BS includes backward- and forward-looking modules. The backward-looking module records realized product-level price changes in the previous calendar month, while the forward-looking module captures firms' expected product-level price changes for the upcoming three calendar months.[10] Survey responses are qualitative in nature —i.e., firms have three reply options: "decrease", "no change", "increase". The qualitative nature is aimed at reducing the reporting burden on the firm, as the survey is not mandatory.

3.2 Measurement

We now describe how we construct the variables for the empirical analysis. Appendix B provides additional details on the data cleaning procedures.

Output prices. The variable of interest is the domestic price charged by firm i, $P_{i,t} \equiv \exp(p_{i,t})$. Constructing this variable requires us to aggregate the product dimension in PRODCOM to a firm-level dimension. Indeed, as shown in Table 1, the mean number of products produced by a single VAT identifier is 3 in our dataset (median 2). To that end, we exploit the fact that the first four digits of this product classification follow the NACE taxonomy, i.e., the official industry nomenclature in the European Union. We define a firm i as the set of products that share the first four NACE digits and are produced by the same VAT identifier. Accordingly, we define an industry as the set of firms whose products share the same first four digits. This leads to 6,223 firms observed over 348 months (1995:m1—2023:m12), resulting in a total of 802,738 observations, and 176 manufacturing industries, distributed across 10 (2-digit) manufacturing sectors. The classification optimally balances a coherent definition of the firm (which is mostly precise if inclusive), with a meaningful

[9] The survey does not require firms to distinguish between production and sales to domestic and international customers. Therefore, we recover domestic values and quantities sold by combining information from PRODCOM with international trade data on firms' product-level exports (quantities and sales).

[10] Firms receive the survey on the first day of the month and are requested to submit their responses within 10 calendar days. The forward-looking module covers the current month and the subsequent two months.

Table 1: Summary statistics

	Mean	p. 5	p. 25	p. 50	p. 75	p. 95
Number of employees	103	9	18	31	66	361
Turnover	2,602	36	164	420	1,195	7,779
Number of products	3	1	1	2	3	9
Number of industries	1	1	1	1	1	2
Revenue share of main industry	98.7	96	100	100	100	100
Number of consecutive months in the sample	120	27	53	96	157	312

Notes. The summary statistics reported in this table refer to the sample of domestic producers in PRODCOM. The sample includes 6,223 firms observed over 348 months (1995:m1–2023:m12), totaling 802,738 observations.

definition of an industry (which is mostly robust if narrow). As we show in Table 1, approximately 90% of the VAT identifiers have products that belong to only one industry, with the main industry accounting for 98.7% of the revenues, on average (median 100%).

Due to repeated product code revisions, a consistent 8-digit product code taxonomy does not exist across the entire sample period. Therefore, we compute the sequence of price changes across consecutive time periods by mapping the product codes at t to their corresponding codes at $t+1$, aggregating them at the firm-industry level, and recovering the time series of the firm-industry price index (in levels) by concatenating monthly price changes. More precisely, denote by $\mathcal{P}_{i,t}$ the set of products manufactured by firm i and by $P_{p,i,t}$ the price (unit value) of a given product $p \in \mathcal{P}_{i,t}$. We first compute the gross price change for each product, $P_{p,i,t}/P_{p,i,t-1}$. In doing so, we appropriately account for any changes in product codes and drop product-level observations with abnormally large price jumps in a given month ($P_{p,i,t}/P_{p,i,t-1} \notin [1/3, 3]$). We then construct the Törnqvist index, which measures the firm-industry price change:

$$P_{i,t}/P_{i,t-1} = \prod_{p \in \mathcal{P}_{i,t}} (P_{p,i,t}/P_{p,i,t-1})^{\bar{s}_{p,i,t}},$$

where $\bar{s}_{p,i,t}$ is a Törnqvist weight computed as the average of the sale shares between t and $t-1$, given by $\bar{s}_{p,i,t} := \frac{s_{p,i,t}+s_{p,i,t-1}}{2}$. The Törnqvist weights, $\bar{s}_{p,i,t}$, give larger weights to those products that account for a larger share of the firm's turnover. We then normalize the first firm-industry observation to unity and concatenate the changes to recover the levels.

Finally, we construct the economy-level price index, used in section 5, as a Törnqvist-

weighted geometric average of price changes, across products and firms:

$$P_t/P_{t-1} = \prod_{\substack{p\in\mathcal{P}_{i,t},\\ i\in[0,1]}} (P_{p,i,t}/P_{p,i,t-1})^{\bar{s}_{p,i,t}}.$$

We normalize the first observation to unity and concatenate the changes to recover levels.

Marginal costs. We construct an empirical measure of marginal cost allowing for general production functions that possibly display non-constant short-run returns to scale. In detail, we relax the assumptions in section 2 and let the marginal cost of firm i depend on the firm's output as $MC^n_{i,t} = \exp(mc^n_t + \xi_{i,t}) \cdot Y^{\nu_i}_{i,t}$, where $1/(1+\nu_i)$ is the firm-specific degree of short-run returns to scale. With constant returns to scale, $\nu_i = 0$. Under this functional form, the nominal marginal cost of a firm is proportional to its average variable costs: $MC^n_{i,t} = (1+\nu_i)AVC_{i,t}$. Using the definition of average variable costs (total variable costs over output, $AVC_{i,t} \equiv TVC^n_{i,t}/Y_{i,t}$) and applying a logarithmic transformation ($avc_{i,t} \equiv \ln(AVC_{i,t})$), we have that firm-level log-nominal marginal cost is given by:

$$mc^n_{i,t} = avc_{i,t} + \ln(1+\nu_i). \tag{20}$$

In the data, we measure total variable costs as the sum of intermediate costs (materials and services purchased) and labor costs (wage bill).[11] We compute a quantity index by dividing a firm's domestic revenue by its domestic price index.[12] We discuss in sections 4.1 and 4.3 how to address concerns regarding classical measurement error and model misspecification due to non-constant returns to scale.

Finally, mirroring the construction of the aggregate price index, we construct the economy-level marginal cost index (used in section 5) as a Törnqvist-weighted geometric average of firm-level marginal cost changes:

$$MC^n_t/MC^n_{t-1} = \prod_{i\in[0,1]} (MC^n_{i,t}/MC^n_{i,t-1})^{\bar{s}_{p,i,t}}.$$

We normalize the first observation to unity and concatenate the changes to recover levels.

[11] While the former is available at the monthly level, the latter is observed at a quarterly frequency and split equally across the three months of the quarter. For multi-industry VAT identifiers, we let the average cost be the same for each industry in which the VAT identifier operates.

[12] We compute $Y_{i,t} = (PY)_{i,t}/\bar{P}_{i,t}$, where $\bar{P}_{i,t}$ denotes the domestic price index for a VAT identifier. For single-industry firms, $\bar{P}_{i,t}$ coincides with the firm-industry price index $P_{i,t}$. For multi-industry firms, $\bar{P}_{i,t}$ is a weighted average of the different firm-industry prices.

Subjective expectations. The NBB-BS survey asks questions regarding product prices at the five-digit PRODCOM level. We use two questions: the *realized* direction of the price change in the current period ($\text{sign}(\Delta p_{i,t})$) and the *expected* direction of the price change in the upcoming three months. With a slight abuse of notation, we denote the latter by $\text{sign}\,\mathbb{E}_{i,t}(\Delta p_{i,t+1})$.[13] We aggregate the responses to the four-digit firm level by averaging. We then linearly interpolate the survey responses when at most one quarter is missing, but both the preceding and following months are present.[14] Finally, we merge the survey, resulting in a final sample of 907 firms, corresponding to 84,699 observations.

Panel A of Table 2 documents that, on average, 79% of firms keep prices fixed compared to the preceding month (consistent with U.S. evidence for producer prices from Nakamura and Steinsson 2008). Similarly, on average, firms expect not to update prices 73% of the time. They expect to decrease prices approximately as often as they do (10% of the times), and expect to increase prices 17% of the times (realized frequency is 11%).

Panel B illustrates how often firms correctly anticipate the direction of a price change. For the vast majority of times, firms' expectations are followed by actions. They correctly expect to decrease prices 52.5% of the times, correctly anticipate to keep prices fixed 82.8% of the times, and increase 63.0% of the times. Sign errors in expectations are rare: a price increase is mistakenly expected to be a decrease in 3.8% of cases, while a price decrease is mistakenly expected to be an increase in 7.8%. This evidence is broadly consistent with both rationality and a Gaussian (thin-tailed) distribution of shocks.

Finally, we show in panel C the forecasting power of expectations for realizations. Since firms report expected price changes over a three-month horizon, we link realized monthly price adjustments to expectations formed one to three months prior. In the first three specifications, we do not include fixed effects and report the intercept, which is not statistically significant. As the mean of $\text{sign}(\Delta p_{i,t})$ is positive (2%), the zero intercept indicates that firms, on average, correctly predict realizations. Furthermore, an expected price change in the previous month predicts a realized price change this month with a probability ranging between 0.4 and 0.26, depending on the specification. Given the low unconditional probabilities of price increases and decreases (9.5% and 11.5%, respectively), these effects are economically substantial. Expectations formed one month prior are the most predictive, while those from two and three months earlier remain significant but markedly smaller, highlighting that near-term expectations carry the most information.

[13] This notation reflects soft information obtained by the data collectors, which indicates that expectations dominantly reflect the nearest term (one to two months). We explore further the effective timing in Table 2.

[14] This imputation permits preserving long time series of consecutive observations for firms. The time dimension of the panel is important because the empirical analysis requires three consecutive periods for constructing the moment condition and several lags of the instruments for estimation. As expectations are only used in the instrument set, these imputations do not introduce potential downward bias from measurement error but ensure consistent results across specifications. See section 4 and Appendix B for further details.

Table 2: NBB Business Survey

Panel A: Summary statistics

	Decrease	Unchanged	Increase	Mean	S.d.
$\text{sign}(\Delta p_{i,t})$	9.49	79.03	11.49	1.99	45.75
$\text{sign}(\mathbb{E}_{i,t-1}(\Delta p_{i,t}))$	9.82	73.00	17.18	7.36	26.45

Panel B: Expectations vs realizations

		$\text{sign}(\Delta p_{i,t})$		
		Decrease	Unchanged	Increase
$\text{sign}(\mathbb{E}_{i,t-1}(\Delta p_{i,t}))$	Decrease	52.54	5.77	3.79
	Unchanged	39.61	82.78	33.20
	Increase	7.85	11.45	63.02

Panel C: Regression analysis

	Dependent variable: sign$(\Delta p_{i,t})$					
$\text{sign}(\mathbb{E}_{i,t-1}(\Delta p_{i,t}))$	0.412	0.314	0.301	0.328	0.271	0.264
	(0.011)	(0.015)	(0.016)	(0.016)	(0.022)	(0.023)
$\text{sign}(\mathbb{E}_{i,t-2}(\Delta p_{i,t-1}))$	–	0.171	0.133	–	0.134	0.110
	–	(0.009)	(0.006)	–	(0.008)	(0.007)
$\text{sign}(\mathbb{E}_{i,t-3}(\Delta p_{i,t-2}))$	–	–	0.078	–	–	0.064
	–	–	(0.007)	–	–	(0.009)
Intercept	-0.009	-0.015	-0.016	–	–	–
	(0.005)	(0.004)	(0.004)	–	–	–
Ind × time FE	×	×	×	✓	✓	✓
Firm FE	×	×	×	✓	✓	✓

Notes. This table refers to the sample of domestic producers in PRODCOM that also participate in the NBB-BS. The subsample includes 907 firms and represents a total of 84,699 observations. $\text{sign}(\cdot)$ variables are encoded as -1 ("decrease"), 0 ("unchanged"), or +1 ("increase"). Panel C reports the results of the regression $\text{sign}(\Delta p_{i,t}) = \alpha + \sum_{h=1}^{3} \gamma_h \text{sign}(\mathbb{E}_{i,t-h}(\Delta p_{i,t-h+1})) + \iota_i + \iota_t + \varepsilon_{i,t}$, where ι_i is a firm fixed effect and ι_t is an industry-by-time (four digits) fixed effect. Standard errors, clustered at the sector level (two digits), are reported in brackets.

4 Empirical analysis

In this section, we discuss the instrument set, estimate the baseline specification, and show the robustness of our results.

4.1 Instrument set

Following the discussion in section 2.5, identification of the structural parameters requires an instrument set that includes shifters of current marginal cost, lagged prices, and expected future prices. Using the administrative panel and the expectations survey, we implement the following instrument set:

$$\mathcal{I}_{i,t}^{IV} = \left[\left\{ avc_{i,t-\tau} \right\}_{\tau=3}^{6}, \{ p_{i,t-\tau} \}_{\tau=2}^{4}, \left\{ \operatorname{sign} \mathbb{E}_{i,t-\tau}(\Delta p_{i,t-\tau+1}) \right\}_{\tau=0}^{1} \right]'.$$

We include lags of average variable cost to address practical concerns such as intra-period production and classical measurement error in contemporaneous cost proxies. Because costs are highly persistent (Table 4), using lagged costs preserves relevance while improving robustness to noise in the contemporaneous measure. We also include the lagged price level to mitigate potential small-sample biases that can arise in dynamic panel settings with fixed effects (Nickell 1981). Finally, we include firms' contemporaneous and lagged subjective expectations of the next-period price change to provide a shifter for future prices that is in the information set at time t and is not spanned by lagged realized prices. As shown in proposition 2, our model implies that these instruments satisfy orthogonality.

4.2 Baseline specification

We perform a two-way demeaning of all the variables and instruments by removing a four-digit industry-by-time fixed effect and a firm fixed effect. Whereas the industry-by-time fixed effect is required by the theory, the firm fixed effect addresses the potential bias from using the average variable cost as a proxy for the marginal cost (equation 20) and improves robustness to potential concerns of permanent heterogeneity (see section 4.3). We allow for correlated shocks across industries in the computation of standard errors. Specifically, we cluster standard errors at the two-digit level so that firms that are in different industries but in the same sector can be exposed to common shocks. We employ a two-step estimator with a weighting matrix at the same level as the clustering (two-digit NACE). Mirroring the construction of the price indexes (section 3.2), all the regressions are weighted using Törnqvist weights. We calibrate the household's discount factor to a standard monthly calibration of $\beta = 0.998$. The ratio of idiosyncratic to total variances is estimated directly

from the data by fitting an AR(1) specification to average cost with and without industry-by-time fixed effects (Table 4) and measuring the variances of the residuals. We obtain $\sigma_\xi^2/\sigma_\varepsilon^2 = 0.96$, consistent with a predominant role of idiosyncratic disturbances. We then estimate specification (19) via the generalized method of moments (Hansen 2010), imposing the coefficient restrictions from theory.[15]

Results are presented in Table 3, column (A). First, we estimate a monthly discount factor of $\omega_f = 0.82$ (0.03)—well below the full-information benchmark of $\beta = 0.998$, and tightly estimated. This estimate implies pronounced myopia: firms incorporate into current prices only a limited share of anticipated cost shocks. Second, we obtain an estimate for the anchoring parameter of $\omega_\ell = 0.10$ (0.02). This coefficient is—perhaps surprisingly—small, indicating little backward-looking persistence in prices at the micro-level beyond what is mechanically generated by nominal rigidity. Third, the estimate for $\omega_p = 1.15$ (0.14) points to a sizeable, though imprecisely estimated, long-run underreaction to idiosyncratic cost shocks, further reducing pass-through relative to the full-information benchmark. These three reduced-form coefficients are jointly disciplined by one overidentifying restriction and two structural parameters, namely the persistence of the belief process, $\lambda = 0.12$ (0.02), and the persistence of the cost process, $\rho = 0.97$ (0.05). The coefficient λ implies a statistically significant departure from the full-information benchmark (which corresponds to $\lambda = 0$); the coefficient ρ is not statistically distinguishable from a direct estimate obtained by fitting an AR(1) process with firm and industry-by-time fixed effects to the average cost (see section 4.3). Fourth, the estimate of price stickiness, $\theta = 0.82$ (0.03), implies infrequent price adjustment—roughly one change every two quarters. This number is remarkably consistent with direct estimates of the monthly fraction of price changes in the NBB-BS data (0.79, Table 2), as well as evidence on the same dataset at a quarterly frequency (Gagliardone et al. 2024). Fifth, we recover a degree of strategic complementarities, $\Omega = 0.52$ (0.06), which is precisely estimated and smaller—though not statistically different— than previous estimates (Amiti et al. 2019, Gagliardone et al. 2025).

These results imply a largely incomplete long-run pass-through of idiosyncratic cost shocks ($(1-\Omega)/\omega_p$), which is approximately 0.42 of that of an aggregate shock. The slope coefficient $\kappa(1-\Omega) = 0.02$ indicates a sizeable contemporaneous pass-through of cost shocks into prices. At the quarterly frequency, this number is given by $\kappa(1-\Omega) = 0.062$, which can be mapped (section 5) into a slope of the cost-based New Keynesian Phillips curve that is an order of magnitude larger than influential estimates in the literature for the output-based curve (Rotemberg and Woodford 1997, Hazell et al. 2022). Finally, the pass-through of an anticipated cost shock in the next month is given by $\kappa(1-\Omega)\omega_f = 0.015$; with

[15] In practice, we estimate the coefficients θ, Ω, ω_f, and ω_p and impose the parameter restrictions on ω_ℓ to improve numerical stability. Appendix A.10 provides the details.

Table 3: Estimates of the parameters

	(A)	**(B)**	**(C)**	**(D)**	**(E)**
	Reduced-form parameters				
ω_f	.821	.824	.803	.873	.667
	(.025)	(.027)	(.088)	(.008)	(.009)
ω_ℓ	.098	.096	.078	.004	0
	(.021)	(.022)	(.024)	(.004)	-
ω_p	1.153	1.165	1.315	1.02	1
	(.147)	(.146)	(.485)	(.034)	-
	Structural parameters				
θ	.827	.831	.809	.8	.8
	(.039)	(.040)	(.032)	-	-
Ω	.516	.519	.421	.55	.55
	(.055)	(.064)	(.176)	-	-
λ	.119	.106	.084	-	-
	(.019)	(.027)	(.028)	-	-
ρ	.967	.925	.816	-	-
	(.056)	(.079)	(.242)	-	-
Firm FE	✓	✓	✓	✓	✓
Ind × time FE	✓	✓	✓	✓	✓
	Sargan-Hansen's overidentification test				
χ^2-stat	6.275	6.193	4.937	7.576	8.827
P-value	.280	.287	.424	.371	.453

Notes. Moments are weighted using Törnqvist weights. GMM is implemented using a two-step estimator with a weighting matrix and standard error clustered at the sector level. Standard errors are reported in brackets.

exponential discounting, the pass-through of an expected shock next quarter is strongly myopic and given by $\kappa(1-\Omega)\omega_f^3 = 0.035$ (using the quarterly slope).

Overall, these estimates support a central role for incomplete information, nominal rigidities, and strategic pricing behavior. The evidence indicates that firms are forward-looking but sharply myopic: they pass through contemporaneous cost shocks strongly, while responding far less to anticipated shocks. Firms therefore behave in a seemingly "static" way despite nominal frictions: they wait for shocks to materialize and then adjust prices aggressively once uncertainty is resolved.

4.3 Robustness analysis

In this section, we check the robustness of the results by exploring alternative assumptions for the model and robustness exercises.

Correcting for returns to scale. One potential concern with the empirical mapping underlying our baseline specification is that the theoretical model treats the real marginal cost, $mc_{i,t}$, as an exogenous process. Yet, as discussed in section 3.2, when firms face non-constant short-run returns to scale (SR-RTS), the nominal marginal cost implied by the production function contains an additional term that varies with output. In logs,

$$mc^n_{i,t} \;=\; mc^n_t \;+\; \xi_{i,t} \;+\; \nu_i\, y_{i,t},$$

where $1/(1+\nu_i)$ measures firm i's SR-RTS and $y_{i,t}$ is the firm's (log) output. Because output is pinned down by demand, $y_{i,t}$ depends endogenously on prices through the demand curve in equation (2). As a result, a cost proxy constructed from observed average variable costs might mechanically inherit an endogenous component.

To address this potential model misspecification, we exploit the sectoral production-function estimates obtained on the same dataset by Gagliardone et al. (2025).[16] Specifically, letting ν_i denote the sector estimate of the SR-RTS parameter of firm i, we purge the output-driven component from average variable cost and construct a corrected measure of the average variable cost:

$$avc^{\text{exo}}_{i,t} \;=\; avc_{i,t} \;-\; \nu_i\, y_{i,t}.$$

We then apply the same demeaning procedure as in the baseline specification and replace the original cost proxy with $avc^{\text{exo}}_{i,t}$ when constructing the moment condition (19) as well as in the instrument set.

Table 3, column (B), reports the resulting estimates. All parameters remain statistically indistinguishable from the baseline, indicating that empirically plausible departures from constant short-run returns to scale do not materially affect our conclusions.

Menu cost models. The empirical specification relies on modeling the nominal rigidity with a constant intensity á la Calvo. This is understood to be a good approximation of menu cost models in "normal times," when the frequency of price adjustments is low and stable, but a worse approximation in unconventional times such as the recent inflation surge (2021-2023).[17] As the latter is the only period of high inflation with large increases

[16] The results, reported for completeness in Appendix B.1, indicate only a small departure from constant returns to scale across sectors. We refer to Gagliardone et al. (2025) for details on the estimation procedure.

[17] See for example Auclert et al. (2024), Blanco et al. (2024), and Gagliardone et al. (2024).

in the frequency in our estimation sample (see Figure A.1 in Appendix B.2), we re-estimate the specification in the pre-Covid period 1995-2018.

The results are reported in Table 3, column (C). The estimates of all the parameters are within the error bands of the baseline specification, with similar point estimates but larger standard errors, likely due to the smaller sample. As most of the variation in the frequency originated from aggregate shocks (e.g., commodity prices, Gagliardone and Gertler 2023), the inclusion of time fixed effects partially controls for the movements in the frequency, leading to consistent conclusions with the baseline analysis.

Test of the overidentifying restriction. A strong assumption underlying the theoretical framework is that of fully rational firms. This assumption implies a one-to-one mapping between the reduced-form coefficients $(\omega_f, \omega_p, \omega_\ell)$ and the structural parameters (σ_u^2, ρ), and therefore delivers a tight overidentifying restriction linking ω_f, ω_p, and ω_ℓ. In practice, departures from full rationality could introduce deviations from this link. We therefore re-estimate the model without imposing the restriction and ask whether the data nonetheless support it. Because the model is not identified in the absence of the restriction, we discipline the estimation by calibrating the remaining two parameters at standard values: we set the Calvo price stickiness to match the frequency of price changes in NBB-BS (Table 2), and fix the degree of strategic complementarities at a benchmark estimate from the literature (Amiti et al. 2019; Gagliardone et al. 2025).

Table 3, column (D), reports the estimates under this alternative specification. The estimated discounting is slightly lower ($\omega_f = 0.87$ (0.01)); the underreaction and anchoring parameters are smaller than the baseline ($\omega_p = 1.02$ (0.03) and $\omega_\ell = 0.004$ (0.004)). This indicates that not imposing the overidentifying restriction tilts the coefficients towards the full-information benchmark, downplaying the role of incomplete information. Notably, in this specification, we find a very limited role for the backward-looking inertia and underreaction, but still a substantial amount of myopia. Finally, we formally test the null hypothesis that the restricted and unrestricted estimates are equal. We implement this test by block bootstrapping with resampling the differences in the estimated coefficients across the two procedures, and conduct a joint test of zero differences.[18] The resulting χ^2-statistic is 0.47 (p-value = 0.92). We therefore fail to reject the null that the data satisfy the overidentifying restriction.

Alternative information structures. The theoretical framework relies on the assumption that firms never directly observe the real marginal cost, but only receive a (possibly perfectly) informative private signal. We now discuss the advantages and limits of this

[18] This approach accounts for the full covariance matrix of the parameters.

modeling choice. First, other firm-level observables such as past firm prices and current or past nominal marginal costs can be included without changing the setup. In this case, the private signal would also account for the information in these additional observables. A caveat is that observing nominal costs should not undo the information friction. Second, including public signals such as the aggregate price level substantially complicates the forecasting step due to the well-known "infinite-regress problem" (Townsend 1983, Nimark 2008, Huo and Takayama 2025), which implies that prices do not generally admit a representation as a finite-order ARMA process. However, our empirical analysis is robust to most of these concerns thanks to the inclusion of time fixed effects, which control for common movements in expectations induced by, e.g., the realization of public signals. Finally, even if firms were to observe current or past real marginal costs, our reduced-form empirical specification could still be motivated in a model in which firms have uncertainty regarding the growth rate of their future costs.

To illustrate the above points formally, consider an alternative information set of the firm given by $\mathcal{I}_{i,t} = \{\{p_{t-\tau}\}_{\tau=0}^{t},\ p_{i,t-1},\ mc_{i,t}^{n}\}$. As in Lucas (1975), at the beginning of each period, firms "pool together" all the information and recover the history of aggregate prices. Additionally, they receive a private signal $s_{i,t} \equiv \{p_{i,t-1}, mc_{i,t}^{n}\}$, from which they infer the real marginal cost, $mc_{i,t}$. Under our maintained assumptions on the primitives, inflation and relative prices are jointly normally distributed with prior distribution given by:

$$\begin{bmatrix} \pi_{t+1} \\ \Delta\widehat{p}_{i,t+1} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbb{E}\left(\pi_{t+1} \;\middle|\; \{p_{t-\tau}\}_{\tau=1}^{t}\right) \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{\pi}^{2} & 0 \\ 0 & \sigma_{\Delta p}^{2} \end{bmatrix} \right)$$

where σ_{π}^{2} and $\sigma_{\Delta p}^{2}$ depend on the variances of aggregate and idiosyncratic shocks. The (posterior) expected growth rate of prices is computed with Bayes rule as:

$$\mathbb{E}_{i,t}(\Delta p_{i,t+1}) = \mathbb{E}\left(\pi_{t+1} \;\middle|\; \{p_{t-\tau}\}_{\tau=0}^{t}\right) + \tilde{\omega}_{f}\, \mathbb{E}\left(\Delta\widehat{p}_{i,t+1} \;\middle|\; s_{i,t}\right)$$

where $\tilde{\omega}_{f} \in [0,1)$ is function of the ratio between the prior precision ($\sigma_{\Delta p}^{-2}$) and the precision of the private signal $s_{i,t}$. Taking differences from the average and combining the above with the pricing equation (12), we obtain:

$$\Delta\widehat{p}_{i,t} = \kappa\left[(1-\Omega)\xi_{i,t} - \widehat{p}_{i,t}\right] + \beta\tilde{\omega}_{f}\, \mathbb{E}\left(\Delta\widehat{p}_{i,t+1} \;\middle|\; s_{i,t}\right) + \psi_{i,t} \tag{21}$$

where $\psi_{i,t} \overset{i.i.d.}{\sim} \mathcal{B}(1-\theta)$ is a Bernoulli mean-zero forecast error with constant intensity $(1-\theta)$. This coincides with the baseline specification (17) under the parameter restrictions $\omega_{\ell} = 0$ and $\omega_{p} = 1$, and relabeling $\omega_{f} \equiv \beta\tilde{\omega}_{f}$.

We estimate specification (21), disciplining estimation as before by restricting the

coefficients of the nominal and real frictions to match the evidence and previous literature. The results are in Table 3, column (E). We estimate a very low discounting (ω_f = 0.66 (0.01)), which is consistent with the main message of the empirical analysis —an essentially static pricing behavior of firms.

Steady-state heterogeneity. The baseline analysis can be generalized with few caveats to accommodate permanent heterogeneity in marginal costs by including a firm-specific drift in equation (3):

$$mc_{i,t+1} = -(1-\rho)\mu_i + \rho mc_{i,t} + \varepsilon_{i,t}.$$

Under this extension, differences across firms driven by, e.g., TFP, network efficiency, or firm appeal may have permanent effects on firms' costs. With steady-state heterogeneity in costs, steady-state sales shares may be heterogeneous as well. In turn, the log-linear markup (equation (9)) would have heterogeneous coefficients (Γ_i) so that the degree of strategic complementarities would depend on the firm's characteristics (Ω_i).

Despite these complications, the empirical results would still be meaningful provided a careful interpretation. Regarding the cost heterogeneity, the firm fixed effects would absorb the firm-specific drift, so that the simple relabeling $\widehat{mc}_{i,t} \equiv mc_{i,t} + \mu_i$ permits the analysis to remain unchanged. Regarding the heterogeneous coefficients, the interpretation of the estimate of the degree of strategic complementarities would change to a Törnqvist-weighted average of heterogeneous firm-specific coefficients. Specifically, our estimates of Ω would recover a local average treatment effect.

Persistence of costs. Our empirical strategy recovers the structural parameters of the model from the estimated reduced-form coefficients, thereby pinning down the persistence of the cost dynamics (ρ) that best matches the dynamics of prices. We can then assess whether this implied persistence is consistent with direct evidence on the evolution of costs, providing an additional validation of both the model assumptions and the empirical results. To this end, we estimate ρ directly from an AR(1) specification for the log average real cost, $avc_{i,t} - p_t$, in three ways: (i) including firm fixed effects, which corresponds to a direct test of the process in equation (3); (ii) including firm and industry-by-time fixed effects, which isolates the persistence of the idiosyncratic component $\xi_{i,t}$; and (iii) aggregating to the economy level, which identifies the persistence of the aggregate component of real marginal cost, mc_t. In all specifications, we instrument $avc_{i,t-1} - p_{t-1}$ with lags two to four to mitigate attenuation bias from measurement error. Results are reported in Table 4. None of the estimates is statistically distinguishable from the others at a 5% level, nor from the structural value in section 4.2 ($\rho = .96$). Therefore, we conclude that the assumption of AR(1) dynamics with equal persistences for the cost is supported in the data.

Table 4: Persistence of the cost components

	$mc_{i,t}$	$mc_{i,t}$	mc_t
ρ	.937	.942	.966
	(.014)	(.011)	(.031)
Firm FE	✓	✓	×
Ind × time FE	×	✓	×

Notes. Estimates of an autoregressive process of order one. The first two specifications have standard errors (in brackets) clustered at the sector level. The third specification has Newey-West robust standard errors (12 lags).

Granularity and demand systems. The theoretical model is derived for the special case of a Kimball aggregator with continuum of monopolistically competitive firms. However, our treatment of the strategic complementarities applies to any model that satisfies, up to first order, the formulation of markups in equation (9). This includes a rather broad class of demand systems with, e.g., homothetic preferences, linear preferences (Melitz and Ottaviano 2008), or translog (Feenstra and Weinstein 2017), as well as alternative market structures such as static or dynamic oligopolies (Atkeson and Burstein 2008, Wang and Werning 2022). A proviso is that, whereas the functional form of output aggregation and competition within a market can often be relaxed, our treatment of the filtering problem would not allow for granular firms manipulating information strategically or internalizing the effects of their choices on other firms' information.

5 Aggregate implications

In section 2, we characterized the "micro" dynamics of relative prices to map the model to the data and recover the structural parameters using only cross-sectional variation (section 4). We now use these estimates to discipline the "macro" dynamics of prices and derive the model's aggregate implications. We first show how to aggregate the economy in the presence of subjective expectations and dispersed information, and we obtain the reduced-form parameters that govern aggregate inflation dynamics by leveraging their dependence on the structural primitives. We then use this calibrated aggregate representation to perform a validation exercise with aggregate data: we construct a time series of real marginal cost from the microdata and feed it into the model to generate the implied inflation path.

5.1 Model aggregation

A key practical obstacle for aggregation in dispersed-information environments is that beliefs are firm-specific. When firms condition on different signals, subjective expectations vary across i, so cross-sectional averaging does not commute with taking expectations: in general, the mean of expectations does not equal the expectation of the mean (e.g., Angeletos and Lian 2018). Our methodology resolves this complexity by returning to the same logic used at the micro level: we project the structural price dynamics onto fundamentals and then impose observational equivalence. This delivers a representation that is both consistent with the incomplete-information equilibrium and easy to aggregate.

Our approach proceeds in two steps. First, we solve for the dynamics of prices under incomplete information. Second, we derive an observationally equivalent complete-information representation, which is then aggregated in closed form. For the first step, we follow similar derivations as those in proposition 1 to characterize the expected inflation rate, $\mathbb{E}_{i,t}(\pi_t)$, whose evolution is described by equation (11) in section 2. As we show in Appendix A.7, the expected inflation rate satisfies an AR(1) process in the signal:

$$\mathbb{E}_{i,t}(\pi_t) = \frac{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)}{(1-\rho\beta)}\left(1-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\frac{1}{(1-\lambda\mathsf{L})}s_{i,t}. \tag{22}$$

Equation (22) coincides with the characterization of inflation in Angeletos and Huo (2021) when there are no idiosyncratic shocks ($\sigma_\xi^2 = 0$). With idiosyncratic shocks, aggregate price growth reflects two distinct components: the evolution of inflation and the evolution of relative prices. In particular, the behavior of $\Delta p_{i,t}$ results from combining the dynamics of relative prices (equation 15) with that of inflation (equation 22). The next proposition shows that the resulting process follows an ARMA(2,1) representation.

Proposition 3 (Dynamics of prices)**.** *The growth rate of prices satisfies the following structural dynamics:*

$$\mathbb{E}_{i,t}(\Delta p_{i,t}) = \underbrace{\frac{\kappa(1-\Omega)(1-\lambda/\rho)}{(\gamma\vartheta-\rho\beta)}}_{\textit{Constant}}\ \underbrace{\frac{(1-\gamma\mathsf{L})}{(1-L/\vartheta)(1-\lambda\mathsf{L})}}_{\textit{ARMA(2,1)}}\ s_{i,t}, \tag{23}$$

where the root of the MA component is

$$\gamma \equiv \frac{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)(1-\rho\beta)+\left(1-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\frac{1}{\vartheta}(\vartheta-\rho\beta)}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)(1-\rho\beta)+\left(1-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)(\vartheta-\rho\beta)} \in \left(\frac{1}{\vartheta},1\right).$$

The proof of the proposition is in Appendix A.7. The key takeaway is that expected price growth blends the relative-price component and the inflation component, with weights governed by the variance ratio $\sigma_\xi^2/\sigma_\varepsilon^2$. Because the relative-price solution already features ARMA(2,1) dynamics (equation 15), the aggregate process inherits the same autoregressive structure. What changes is the moving-average component: the MA root is generally non-trivial ($\gamma < 1$), except in the limiting case with no aggregate shocks ($\sigma_\xi^2/\sigma_\varepsilon^2 \to 1$), in which the MA part becomes degenerate. At the opposite limit, when there are only aggregate shocks ($\sigma_\xi^2/\sigma_\varepsilon^2 \to 0$), the MA root satisfies $\gamma \to 1/\vartheta$ and simplifies with one of the AR roots, so that equation (23) collapses to equation (22). Away from these cases, the dynamics of prices is hump-shaped, with shocks affecting prices persistently both directly and endogenously via the update in beliefs.

The second step to aggregation mirrors the construction used to derive the moment condition in proposition 2. We project the expected dynamics implied by the incomplete-information solution onto the history of observable fundamentals and require that this projection preserves the equilibrium dynamics. Formally, letting $h_a^{\text{CI}}(\mathsf{L})$ denote the resulting reduced-form projection, we impose the following observational equivalence restriction:

$$\mathbb{E}\Big(\mathbb{E}_{i,t}(\Delta p_{i,t}) \,\Big|\, \{mc_{i,\tau}\}_{\tau \le t} \Big) \;=\; h_a^{\text{CI}}(\mathsf{L})\; mc_{i,t} \tag{24}$$

Equation (24) delivers a complete-information benchmark in which $\Delta p_{i,t}$ can be expressed as a function of real marginal cost and lagged price growth rates, but with reduced-form coefficients that are distorted to replicate the effects of dispersed information. Once written in this form, aggregation becomes straightforward: integrating $h_a^{\text{CI}}(\mathsf{L})\; mc_{i,t}$ across firms yields an aggregate Phillips curve. The next proposition provides the resulting aggregate inflation dynamics and highlights how information shapes the reduced-form parameters.

Proposition 4 (Phillips curve)**.** *Inflation satisfies the following reduced-form dynamics:*

$$\pi_t = \kappa(1-\Omega)mc_t + \omega_c^a mc_{t-1} + \omega_f^a \mathbb{E}_t\{\pi_{t+1}\} + \omega_{\ell,1}^a \pi_{t-1} + \omega_{\ell,2}^a \pi_{t-2}, \tag{25}$$

where the coefficients are uniquely pinned down by the structural parameters, as follows:

$$\omega_f^a = \frac{\rho + \beta\rho^2 - \gamma\rho\vartheta - \lambda}{(\rho-\lambda)(\rho+\lambda+1/\vartheta-\gamma)}$$

$$\omega_c^a = -\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)\frac{\gamma}{\gamma\vartheta-\rho\beta}\left(1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)\right)$$

$$\omega_{\ell,1}^a = \left(\lambda+\frac{1}{\vartheta}\right)\left(1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)\right)+\omega_f^a\left(\frac{\lambda}{\vartheta}\right)$$

$$\omega_{\ell,2}^a = -\left(\frac{\lambda}{\vartheta}\right)\left(1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)\right).$$

When $\rho+\lambda+1/\vartheta > \gamma$, the coefficients satisfy $\omega_f^a < \beta$, $\omega_c^a < 0$, $\omega_{\ell,1}^a > 0$, and $\omega_{\ell,2}^a < 0$.

The proof of the proposition is in Appendix A.8. The Phillips curve in (25) takes a "hybrid" form (Galí and Gertler 1999). As in the standard New Keynesian benchmark, inflation is a forward-looking function of real aggregate marginal cost, with slope $\kappa(1-\Omega)$. Our model, however, displays additional features. First, the aggregate discounting, ω_f^a, is smaller than the household's discounting β at a realistic calibration of the model ($\rho+\lambda+1/\vartheta > \gamma$). Second, inflation is partially anchored to past inflation through π_{t-1} and π_{t-2}, governed by $\omega_{\ell,1}^a$ and $\omega_{\ell,2}^a$. Third, inflation inherits a moving-average component through lagged marginal cost, reflected in ω_c^a. These distortions are necessary to match the hump-shaped response of inflation to cost shocks. As discussed in section 2.2, the information friction does not completely vanish when the signal becomes perfectly informative ($\sigma_\varepsilon^2/\sigma_u^2 \to \infty$). As firms cannot dynamically disentangle the two components of real marginal cost, and expected relative prices display a different dynamics compared to expected inflation (equations 11 and 12), the reduced-form aggregate parameters reflect both the signal noise and the interaction between idiosyncratic and aggregate cost shocks. As a result, inflation dynamics is more than just the sum of its parts.

Finally, we connect the "macro" discounting to the "micro" discounting from section 2. The following corollary shows that, when idiosyncratic shocks are the main driver of firm's costs, the aggregate discounting is bounded above by our micro-estimate.

Corollary 3 (Macro vs micro discounting)**.** *There exists a unique threshold $r^\star \in (0,1]$ such that:*

$$\omega_f^a < \omega_f \quad \Longleftrightarrow \quad \frac{\sigma_\xi^2}{\sigma_\varepsilon^2} > r^\star.$$

Table 5: Estimates of the Phillips curve

	(A)	(B)	(C)
ω_f^a	.804	.807	.761
	(.026)	(.031)	(.099)
ω_c^a	-.015	-.014	-.018
	(.012)	(.012)	(.007)
$\omega_{\ell,1}^a$	.304	.299	.338
	(.038)	(.043)	(.063)
$\omega_{\ell,2}^a$	-.022	-.022	-.022
	(.006)	(.007)	(.007)

Notes. The coefficients and standard errors (in brackets) are obtained using the formulas in proposition 4 and the estimates in Table 3. Moments are weighted using Törnqvist weights. GMM is implemented using a two-step estimator with a weighting matrix and standard error clustered at the sector level.

The proof of the corollary is in Appendix A.9. The corollary highlights that the mapping from micro to macro discounting is not mechanical: aggregation can amplify the apparent degree of myopia. When idiosyncratic disturbances account for a sufficiently large share of price volatility (i.e., when $\sigma_\xi^2/\sigma_\varepsilon^2$ exceeds the threshold $r^\star$), the aggregate discount factor ω_f^a is smaller than the micro parameter ω_f estimated in section 4. Combined with the micro estimates and the evidence that $\sigma_\xi^2/\sigma_\varepsilon^2$ is large, this implies a substantial amount of discounting at the aggregate level.

5.2 Estimates of the Phillips curve

The main advantage of our aggregation method is that it provides a one-to-one mapping from the "micro-estimates" in section 4 obtained using only cross-sectional variation to the "macro-coefficients" that regulate the dynamics of inflation in proposition 4. As this approach requires imposing the overidentifying restriction, we can only construct the estimates corresponding to the baseline specification (section 4.2) and the two robustness exercises to menu cost and decreasing returns to scale (section 4.3).

Columns (A)-(C) in Table 5 display the estimates associated, respectively, with columns (A)-(C) of Table 3. We recover a large amount of aggregate discounting in all specifications, with a baseline estimate of $\omega_f^a = 0.8$ (0.02). This number is smaller, though within one standard deviation of the micro-discounting, given by $\omega_f = 0.82$. Therefore, the high discounting at the micro-level translates into a large discounting at the macro-level. As argued in Gabaix (2020), such high myopia can help rationalize the well-known "forward-

guidance puzzle," which has been discussed by Del Negro et al. (2023) and Bocola et al. (2024) among others. On the other hand, whereas relative prices display a low degree of anchoring to past decisions ($\omega_\ell = 0.1$), the model implies that inflation is much more persistent, with a benchmark coefficient on the first lag that is three times larger and precisely estimated ($\omega^a_{\ell,1} = 0.3$ (0.04)). We highlight how these two findings of high discounting and anchoring at the aggregate level implied by the micro-estimates are consistent with those of Galí and Gertler (1999) obtained from time-series aggregate data. Finally, we find quantitatively small coefficients on the moving average component ($\omega^a_c = -0.02$ (0.01)) and the second lag of inflation ($\omega^a_{\ell,2} = -0.02$ (0.01)), suggesting that also a parsimonious representation of the Phillips curve with myopia and only one lag of inflation can provide an adequate fit of the empirical evidence.

5.3 How much inflation volatility can the model explain?

Using data on costs, we now assess the ability of the model to explain inflation dynamics. This exercise serves as a validation based on aggregate time-series data of both the reduced-form dynamics for inflation derived in this section, as well as the micro-estimates obtained from cross-sectional variation and their mapping into the aggregate coefficients.

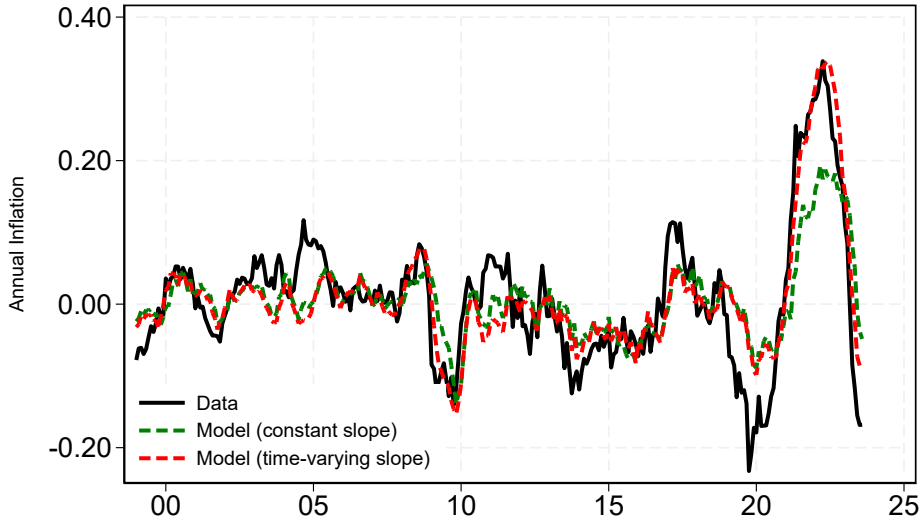
Specifically, solving backward the Phillips curve in equation (25), we derive the model-based relation between year-over-year inflation and the path of aggregate nominal marginal cost. As shown in Appendix A.11, this takes the following form:

$$p_t - p_{t-12} = \sum_{\tau=0}^{11} \left(\Psi_{c,\tau} mc^n_{t-\tau} + \Psi_{\ell,\tau} mc^n_{t-\tau-1} \right) + \sum_{\tau=0}^{2} \Psi_{p,\tau} p_{t-12-\tau} \tag{26}$$

where the coefficients $\Psi_{c,\tau}, \Psi_{\ell,\tau}, \Psi_{p,\tau}$ depend on the structural parameters that we estimated in section 4. We calibrate these at the benchmark specification, column (A) of Table 3. We then aggregate the firm-level nominal marginal costs into an aggregate index, mc^n_t, as a Törnqvist-weighted average and feed it into the model via equation (26).

This exercise produces a model-based series for year-over-year inflation as a function of the price level in the previous year (and two lags of it) and the path of realized aggregate nominal marginal cost. We confront this series against the year-over-year inflation in the data. Figure 1 shows the results. The black solid line is the data, the green dashed line is the model-implied inflation. The model tracks inflation remarkably well. In particular, in the pre-Covid period, the model aligns with the data both in terms of realized volatility and cyclical movements, indicating a dominant role of cost-driven fluctuations for inflation. We regress the data on the model-based series and obtain an $R^2 = 0.63$, which indicates that approximately two-thirds of the variation in inflation is attributable to cost shocks.

Figure 1: Inflation (year-over-year) in data and model



Notes. Time series of year-over-year aggregate inflation and model-based year-over-year inflation. All variables are aggregated using Törnqvist weights. The aggregate price level and aggregate nominal marginal cost are detrended by removing a piece-wise linear trend with one break. The model-based inflation series are smoothed using a centered moving average with two leads and lags. The green dashed line is the model-based year-over-year inflation series (equation (26)). The red dashed line is the model-based year-over-year series calibrated to match the aggregate time-varying frequency of price changes in the data. Data is the black solid line.

The baseline model also performs well during the Covid pandemic and recent inflation surge, though it does not fully capture the amplitude of the fluctuations observed in the data. This is consistent with a potential central role for demand-driven inflation and non-linearities in the slope of the Phillips curve, two of the leading themes discussed in the current macroeconomic debate. To inform about their importance, we recalibrate the coefficients of equation (26) to allow for a time-varying slope of the Phillips curve that matches the realized frequency of price changes (Figure A.1 in Appendix B.2). The model-based series is reported as the red dashed line of Figure 1. The model does substantially better in capturing the inflation surge, but does not improve much compared to the baseline during the remaining part of the sample, including the Covid period. This evidence suggests a key role for demand during the pandemic, and a central role for non-linearities in the inflation surge that followed. As before, we regress the data on the model-based series, obtaining an OLS estimate of 0.98 and an $R^2 = 0.73$. This reinforces our conclusion that the model can account extremely well for the dynamics of inflation.

6 Conclusions

Using microdata on firms' costs, prices, and expectations, we demonstrated how their pricing behavior is consistent with a high discounting of anticipated future cost shocks. Our benchmark estimate of the firm's effective discount factor is roughly 0.8 at a monthly frequency, well below the standard calibration of the household's discount factor of approximately one. We rationalized this evidence through the lense of a model that combines incomplete information with nominal and real frictions, and delivers a mapping from structural parameters to a reduced-form empirical specification with an overidentifying restriction. Even if, in the model, firms are rational and use information optimally, incomplete information distorts their pricing behavior, making it appear as if they are myopic.

This evidence provides a novel perspective on how firms set prices at the micro-level and their aggregate implications. Our micro-estimates are consistent with an environment in which firms set prices almost statically: they wait for cost shocks to materialize before adjusting prices. Then, as costs change, firms absorb only a fraction of the shock into markups while passing a significant share into prices. These findings have relevant implications also for the debate on the flattening of the slope of the Phillips curve. At the aggregate level, our model implies a quarterly slope of approximately 6%, pronounced aggregate myopia (high aggregate discounting), and anchoring to past inflation (backward-looking persistence). The weak response of aggregate prices to highly persistent real shocks that has been observed in the first two decades of the century can be rationalized by the model. Firms do not respond to the distant path of future shocks and inflation is anchored to the past, resulting in an estimated flat slope when the discounting is calibrated to conventional values. Finally, we show that variation in costs can explain two-thirds of inflation volatility, indicating a key role for supply-driven inflation.

An interesting venue for future research is the role of heterogeneity and time variation in the model's parameters. Changes in the slope of the Phillips curve due to increases in the measured frequency of price changes are already understood to be instrumental to explaining the recent inflation surge (Gagliardone et al. 2024). Complementarily, our theory shows that the firms' effective discounting is not a structural parameter and, therefore, is likely to change endogenously as a function of the information available to firms. As periods in which inflation is high are typically associated with more attention towards aggregate variables (Pfäuti 2023, Weber et al. 2025), cyclical fluctuations in the discounting may further help to bridge the gap between data and theory on inflation dynamics.

References

Hassan Afrouzi, Joel P Flynn, and Choongryul Yang. What can measured beliefs tell us about monetary non-neutrality? Technical report, National Bureau of Economic Research, 2024.

Mary Amiti, Oleg Itskhoki, and Jozef Konings. International shocks, variable markups, and domestic prices. *The Review of Economic Studies*, 86(6):2356–2402, 2019.

Mary Amiti, Oleg Itskhoki, and Jozef Konings. Dominant currencies: How firms choose currency invoicing and why it matters. *The Quarterly Journal of Economics*, 137(3):1435–1493, 2022.

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

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

George-Marios Angeletos and Jennifer La'O. Incomplete information, higher-order beliefs and price inertia. *Journal of Monetary Economics*, 56:S19–S37, 2009.

George-Marios Angeletos and Jennifer La'O. Optimal monetary policy with informational frictions. *Journal of Political Economy*, 128(3):1027–1064, 2020.

George-Marios Angeletos and Chen Lian. Forward guidance without common knowledge. *American Economic Review*, 108(9):2477–2512, 2018.

Andrew Atkeson and Ariel Burstein. Pricing-to-market, trade costs, and international relative prices. *American Economic Review*, 98(5):1998–2031, 2008.

Adrien Auclert, Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub. New pricing models, same old phillips curves? *The Quarterly Journal of Economics*, 139(1):121–186, 2024.

Raphael A Auer and Raphael S Schoenle. Market structure and exchange rate pass-through. *Journal of International Economics*, 98:60–77, 2016.

Isaac Baley and Andrés Blanco. Firm uncertainty cycles and the propagation of nominal shocks. *American Economic Journal: Macroeconomics*, 11(1):276–337, 2019.

Andres Blanco, Corina Boar, Callum J Jones, and Virgiliu Midrigan. Non-linear inflation dynamics in menu cost economies. Technical report, National Bureau of Economic Research, 2024.

Luigi Bocola, Alessandro Dovis, Kasper Jørgensen, and Rishabh Kirpalani. Bond market views of the fed. Technical report, National Bureau of Economic Research, 2024.

Pedro Bordalo, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. Overreaction in macroeconomic expectations. *American Economic Review*, 110(9):2748–2782, 2020.

Guillermo A Calvo. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12(3):383–398, 1983.

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

Olivier Coibion, Yuriy Gorodnichenko, and Saten Kumar. How do firms form their expectations? New survey evidence. *American Economic Review*, 108(9):2671–2713, 2018.

Olivier Coibion, Yuriy Gorodnichenko, and Tiziano Ropele. Inflation expectations and firm decisions: New causal evidence. *The Quarterly Journal of Economics*, 135(1):165–219, 2020.

Isabelle De Greef and Christophe Van Nieuwenhuyze. The national bank of belgium's new business survey indicator. *NBB Economic Review*, June:31–51, 2009.

Marco Del Negro, Marc P Giannoni, and Christina Patterson. The forward guidance puzzle. *Journal of Political Economy: Macroeconomics*, 1(1):43–79, 2023.

Keshav Dogra, Sebastian Heise, Edward S Knotek, Brent Meyer, Robert W Rich, Raphael Schoenle, Giorgio Topa, Wilbert Van der Klaauw, and Wändi Bruine de Bruin. Estimates of cost-price passthrough from business survey data. 2023.

Robert F Engle and Clive WJ Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

European Commission. Quality report on prodcom 2014 annual data. PRODCOM Working Group, 2014.

Emmanuel Farhi and Iván Werning. Monetary policy, bounded rationality, and incomplete markets. *American Economic Review*, 109(11):3887–3928, 2019.

Robert C Feenstra and David E Weinstein. Globalization, markups, and us welfare. *Journal of Political Economy*, 125(4):1040–1074, 2017.

Xavier Gabaix. A behavioral new keynesian model. *American Economic Review*, 110(8): 2271–2327, 2020.

Xavier Gabaix and David Laibson. Myopia and discounting. Technical report, National bureau of economic research, 2017.

Luca Gagliardone and Mark Gertler. Oil prices, monetary policy and inflation surges. Technical report, National Bureau of Economic Research, 2023.

Luca Gagliardone, Mark Gertler, Simone Lenzu, and Joris Tielens. Micro and macro cost-price dynamics across inflation regimes. Technical report, 2024.

Luca Gagliardone, Mark Gertler, Simone Lenzu, and Joris Tielens. Anatomy of the phillips curve: micro evidence and macro implications. *American Economic Review*, 115(11): 3941–3974, 2025.

Jordi Galí and Mark Gertler. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics*, 44(2):195–222, 1999.

Mishel Ghassibe, Boromeus Wanengkirtyo, and Ivan Yotzov. Intertemporal pass-through. 2025.

Pinelopi Koujianou Goldberg and Rebecca Hellerstein. A structural approach to explaining incomplete exchange-rate pass-through and pricing-to-market. *American Economic Review*, 98(2):423–429, 2008.

Lars Peter Hansen. Generalized method of moments estimation. *Macroeconometrics and time series analysis*, pages 105–118, 2010.

Jonathon Hazell, Juan Herreno, Emi Nakamura, and Jón Steinsson. The slope of the phillips curve: evidence from us states. *The Quarterly Journal of Economics*, 137(3):1299–1344, 2022.

Christian Hellwig and Venky Venkateswaran. Dispersed information, sticky prices and monetary business cycles: A hayekian perspective. *New York University, Mimeo*, 2014.

Zhen Huo and Naoki Takayama. Rational expectations models with higher-order beliefs. *Review of Economic Studies*, 92(5):3138–3173, 2025.

Miles S Kimball. The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit, and Banking*, 27(4):1241–1277, 1995.

Simone Lenzu, David Rivers, Joris Tielens, and Hu Shi. Financial shocks, productivity, and prices. *Review of Economic Studies, forthcoming*, 2025.

Jr Robert E Lucas. Expectations and the neutrality of money. *Journal of Economic Theory*, 4 (2):103–124, 1972.

Robert E Lucas. An equilibrium model of the business cycle. *Journal of Political Economy*, 83(6):1113–1144, 1975.

Bartosz Maćkowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009.

N Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4): 1295–1328, 2002.

Michael McLeay and Silvana Tenreyro. Optimal inflation and the identification of the phillips curve. *NBER Macroeconomics Annual*, 34(1):199–255, 2020.

Marc J Melitz and Gianmarco IP Ottaviano. Market size, trade, and productivity. *The Review of Economic Studies*, 75(1):295–316, 2008.

Leonardo Melosi. Estimating models with dispersed information. *American Economic Journal: Macroeconomics*, 6(1):1–31, 2014.

Emi Nakamura and Jón Steinsson. Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4):1415–1464, 2008.

Emi Nakamura and Dawit Zerom. Accounting for incomplete pass-through. *The Review of Economic Studies*, 77(3):1192–1230, 2010.

Stephen Nickell. Biases in dynamic models with fixed effects. *Econometrica*, 49(6):1417–1426, 1981.

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

Luigi Paciello and Mirko Wiederholt. Exogenous information, endogenous information, and optimal monetary policy. *Review of Economic Studies*, 81(1):356–388, 2014.

Oliver Pfäuti. The inflation attention threshold and inflation surges. *arXiv preprint arXiv:2308.09480*, 2023.

Julio J Rotemberg and Michael Woodford. An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual*, 12:297–346, 1997.

Robert M Townsend. Forecasting the forecasts of others. *Journal of Political Economy*, 91 (4):546–588, 1983.

Olivier Wang and Iván Werning. Dynamic oligopoly and price stickiness. *American Economic Review*, 112(8):2815–49, 2022.

Michael Weber, Bernardo Candia, Hassan Afrouzi, Tiziano Ropele, Rodrigo Lluberas, Serafin Frache, Brent Meyer, Saten Kumar, Yuriy Gorodnichenko, Dimitris Georgarakos, et al. Tell me something i don't already know: Learning in low-and high-inflation settings. *Econometrica*, 93(1):229–264, 2025.

Michael Woodford. Imperfect common knowledge and the effects of monetary policy, 2001.

Dynamic Pricing under Information Frictions: Evidence from Firm-level Subjective Expectations

L. Gagliardone J. Tielens

Appendix

A Proofs for the text

This section includes the derivations of the equations and proofs of propositions.

A.1 Derivations of equation (9)

Assume that output Y_t is produced by a unitary measure of perfectly competitive firms using a bundle of differentiated intermediate inputs $Y_{i,t}$. The bundle of inputs is assembled into final goods using the Kimball aggregator:

$$\int_0^1 \Upsilon\left(\frac{Y_{i,t}}{Y_t}\right) di = 1,$$

where $\Upsilon(\cdot)$ is strictly increasing, strictly concave, and satisfies $\Upsilon(1) = 1$. Taking as given the demand Y_t, each firm minimizes costs subject to the aggregate constraint:

$$\min_{Y_{i,t}} \int_0^1 P_{i,t} Y_{i,t} di \quad \text{s.t.} \int_0^1 \Upsilon\left(\frac{Y_{i,t}}{Y_t}\right) di = 1.$$

Denoting by ψ the Lagrange multiplier of the constraint, the first-order condition of the problem is:

$$P_{i,t} = \psi \Upsilon'\left(\frac{Y_{i,t}}{Y_t}\right) \frac{1}{Y_t} \tag{A.1}$$

Define implicitly the aggregate price index P_t as:

$$\int_0^1 \phi\left(\Upsilon'(1)\frac{P_{i,t}}{P_t}\right) di = 1$$

where $\phi \equiv \Upsilon \circ (\Upsilon')^{-1}$. Evaluating the first-order condition (A.1) at symmetric prices, $P_{i,t} = P_t$, we get $\psi = \frac{P_t Y_t}{\Upsilon'(1)}$. Replacing for ψ, we recover the demand function:

$$\frac{P_{i,t}}{P_t} = \frac{1}{\Upsilon'(1)} \Upsilon' \left(\frac{Y_{i,t}}{Y_t} \right). \tag{A.2}$$

Therefore, the demand function faced by firms when resetting prices is:

$$Y_{i,t} = \left[(\Upsilon')^{-1} \left(\Upsilon'(1) \frac{P^o_{i,t}}{P_t} \right) \right] Y_t$$

Taking logs of equation (A.1) and differentiating, we obtain the following expression for the residual elasticity of demand:

$$\epsilon_{fit} \equiv -\frac{\partial \ln Y_{i,t}}{\partial \ln P^o_{i,t}} = -\frac{\Upsilon' \left(\frac{Y_{i,t}}{Y_t} \right)}{\Upsilon'' \left(\frac{Y_{i,t}}{Y_t} \right) \cdot \left(\frac{Y_{i,t}}{Y_t} \right)} \tag{A.3}$$

Denote the steady-state residual demand elasticity by $\epsilon = -\frac{\Upsilon'(1)}{\Upsilon''(1)}$. Then the derivative of the residual demand elasticity $\epsilon_{i,t}$ in (A.3) with respect to $\frac{Y_{i,t}}{Y_t}$, evaluated at the steady state, is given by:

$$\epsilon' = \frac{\Upsilon'(1)(\Upsilon'''(1) + \Upsilon''(1)) - (\Upsilon''(1))^2}{(\Upsilon''(1))^2} \leq 0. \tag{A.4}$$

The equation above holds with equality if the elasticity is constant (e.g., under CES preferences). Log-linearizing the Lerner index around the steady state and using equation (A.4), we have that, up to a first-order approximation, the log-markup (in deviation from the symmetric steady state) is equal to:

$$\mu_{i,t} - \mu = \frac{\epsilon'}{\epsilon(\epsilon - 1)} (y_{i,t} - y_t)$$

Finally, log-linearizing the demand function (A.1), we obtain:

$$p^o_{i,t} - p_t = \frac{\Upsilon''(1)}{\Upsilon'(1)} (y_{i,t} - y_t)$$

and use it to replace the log difference in output:

$$\mu_{i,t} - \mu = -\Gamma \left(p^o_{i,t} - p_t \right)$$

where the sensitivity of the markup to the relative price is given by $\Gamma \equiv -\frac{\epsilon'}{\epsilon(\epsilon-1)} \frac{\Upsilon'(1)}{\Upsilon''(1)} \geq 0$. In the text, we set $\mu = 0$ from the normalization of the steady-state real marginal cost.

A.2 Derivation of equation (12)

Starting from equation (10), take expectations and subtract $p_{i,t-1}$ from both sides to obtain:

$$\mathbb{E}_{i,t}(p^o_{i,t}-p_{i,t-1}) \;=\; \mathbb{E}_{i,t}\Big\{(1-\beta\theta)\Big[(1-\Omega)(\mu+mc^n_{i,t}-p_{i,t-1})+\Omega(p_t-p_{i,t-1})\Big]+\beta\theta(p^o_{i,t+1}-p_{i,t}+\Delta p_{i,t})\Big\}$$

Rewriting in terms of real marginal cost $mc_{i,t} \equiv mc^n_{i,t} - p_t$:

$$\mathbb{E}_{i,t}(p^o_{i,t}-p_{i,t-1}) \;=\; \mathbb{E}_{i,t}\Big\{(1-\beta\theta)\Big[(1-\Omega)(\mu+mc_{i,t})+(p_t-p_{i,t})+\Delta p_{i,t}\Big]+\beta\theta(p^o_{i,t+1}-p_{i,t}+\Delta p_{i,t})\Big\}$$

Multiply both sides by $(1-\theta)$:

$$(1-\theta)\,\mathbb{E}_{i,t}(p^o_{i,t}-p_{i,t-1}) \;=\; \mathbb{E}_{i,t}\Big\{(1-\beta\theta)(1-\theta)\Big[(1-\Omega)(\mu+mc_{i,t})+(p_t-p_{i,t})+\Delta p_{i,t}\Big]+\beta\theta(1-\theta)(p^o_{i,t+1}-p_{i,t}+\Delta p_{i,t})\Big\}$$

Taking expectations of the law of motion (5),

$$\mathbb{E}_{i,t}(\Delta p_{i,t}) = (1-\theta)\,\mathbb{E}_{i,t}(p^o_{i,t}-p_{i,t-1})$$

and

$$\mathbb{E}_{i,t}(\Delta p_{i,t+1}) = (1-\theta)\,\mathbb{E}_{i,t}(p^o_{i,t+1}-p_{i,t})$$

Replace in the above:

$$\mathbb{E}_{i,t}(\Delta p_{i,t}) \;=\; \mathbb{E}_{i,t}\Big\{(1-\beta\theta)(1-\theta)\Big[(1-\Omega)(\mu+mc_{i,t})+(p_t-p_{i,t})+\Delta p_{i,t}\Big]+\beta\theta\Delta p_{i,t+1}+\beta\theta(1-\theta)\Delta p_{i,t}\Big\}$$

Collecting $\Delta p_{i,t}$ on the LHS and dividing by θ:

$$\mathbb{E}_{i,t}(\Delta p_{i,t}) \;=\; \mathbb{E}_{i,t}\Big\{\kappa\Big[(1-\Omega)(\mu+mc_{i,t})+(p_t-p_{i,t})\Big]+\beta\Delta p_{i,t+1}\Big\}$$

Adding and subtracting the cross-sectional means inside the expectations, the above can be decomposed additively as follows:

$$\mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t}) \;=\; \mathbb{E}_{i,t}\Big\{\kappa\Big[(1-\Omega)\xi_{i,t}-\widehat{p}_{i,t}\Big]+\beta\Delta\widehat{p}_{i,t+1}\Big\}$$

$$\mathbb{E}_{i,t}(\pi_t) \;=\; \mathbb{E}_{i,t}\Big\{\kappa(1-\Omega)(\mu+mc_t)+\beta\pi_{t+1}\Big\}$$

where $\mathbb{E}_{i,t}(\Delta p_{i,t}) = \mathbb{E}_{i,t}(\Delta\widehat{p}_{i,t}) + \mathbb{E}_{i,t}(\pi_t)$ and $\mathbb{E}_{i,t}(mc_{i,t}) = \mathbb{E}_{i,t}(\xi_{i,t}) + \mathbb{E}_{i,t}(mc_t)$. In the text, we set $\mu = 0$ from the normalization of the steady-state real marginal cost.

A.3 Kalman filter

State-space representation. The evolution of the state is:

$$mc_{i,t} = \underbrace{\rho}_{F} mc_{i,t-1} + \underbrace{\begin{bmatrix}\sigma_\varepsilon & 0\end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix}\varepsilon_{i,t}\\ u_{i,t}\end{bmatrix}}_{\eta_{i,t}}$$

The measurement equation is:

$$s_{i,t} = \underbrace{1}_{H} mc_{i,t} + \underbrace{\begin{bmatrix}0 & \sigma_u\end{bmatrix}}_{\Psi} \begin{bmatrix}\varepsilon_{i,t}\\ u_{i,t}\end{bmatrix}.$$

Riccati equation and Kalman gain. Let $P \equiv \mathrm{Var}(mc_{i,t} \mid s_{1:t-1})$ be the prior variance. The steady-state Kalman gain is:

$$K \equiv P H^\top (H P H^\top + \sigma_u^2)^{-1} = \frac{P}{P + \sigma_u^2}$$

The Riccati equation is:

$$P \;=\; F[P - KHP]F^\top + \Phi\Phi^\top = \rho^2(1-K)P + \sigma_\varepsilon^2.$$

Define

$$\lambda \equiv \rho(1-K).$$

Plugging into the Riccati and solving:

$$P = \frac{\sigma_\varepsilon^2}{1-\rho\lambda}$$

Plugging into the definition of the Kalman gain:

$$P = \sigma_u^2\left(\frac{\rho}{\lambda} - 1\right) = \sigma_u^2 \frac{\rho - \lambda}{\lambda}.$$

Equating the two equations above leads to a quadratic equation for λ:

$$\lambda^2 - \left[\rho + \frac{1}{\rho}\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_u^2}\right)\right]\lambda + 1 = 0.$$

The stabilizing solution that satisfies $0 < \lambda < \rho$ is:

$$\lambda = \frac{1}{2}\left[\rho + \frac{1}{\rho}\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_u^2}\right) - \sqrt{\left(\rho + \frac{1}{\rho}\left(1 + \frac{\sigma_\varepsilon^2}{\sigma_u^2}\right)\right)^2 - 4}\right]$$

Closed-loop for beliefs. The expectation of the state is:

$$\mathbb{E}_{i,t}(mc_{i,t}) \;=\; (I - KH)F\,\mathbb{E}_{i,t-1}(mc_{i,t-1}) + K\,s_{i,t} = \lambda\,\mathbb{E}_{i,t-1}(mc_{i,t-1}) + \left(1 - \frac{\lambda}{\rho}\right)s_{i,t}$$

Using the modified signal $\widehat{s}_{i,t} \equiv \frac{\sigma_\xi^2}{\sigma_\varepsilon^2} s_{i,t}$, the evolution of the expected idiosyncratic state is:

$$\mathbb{E}_{i,t}(\xi_{i,t}) \;=\; (I - (I - KH)F\mathsf{L})^{-1} K\,\widehat{s}_{i,t} = \frac{1}{1 - \lambda\mathsf{L}}\left(1 - \frac{\lambda}{\rho}\right)\widehat{s}_{i,t}$$

where L denotes the lag operator. The evolution of the expected aggregate state is:

$$\mathbb{E}_{i,t}(mc_t) \;=\; = \frac{1}{1 - \lambda\mathsf{L}}\left(1 - \frac{\lambda}{\rho}\right)\frac{\sigma_a^2}{\sigma_\varepsilon^2} s_{i,t}$$

Fundamental representation. Solving, the state as a function of shocks is:

$$mc_{i,t} = (I - F\mathsf{L})^{-1}\Phi\,\eta_{i,t} = \left[\tfrac{\sigma_\varepsilon}{1-\rho\mathsf{L}} \quad 0\right]\eta_{i,t},$$

and the signal as a function of shocks is:

$$s_{i,t} = \underbrace{\left(H(I - F\mathsf{L})^{-1}\Phi + \Psi\right)}_{=:M(\mathsf{L})}\eta_{i,t} = \left[\tfrac{\sigma_\varepsilon}{1-\rho\mathsf{L}} \quad \sigma_u\right]\eta_{i,t}$$

Hence, the modified signal can be written as:

$$\widehat{s}_{i,t} = (M(\mathsf{L}))^{-1}\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\eta_{i,t}$$

Forecast error. The forecast error is given by:

$$\begin{aligned} w_{i,t} &\equiv s_{i,t} - \mathbb{E}_{i,t-1}(s_{i,t}) \\ &= \underbrace{(I - H(I - F(I - KH)\mathsf{L})^{-1} FK\mathsf{L})}_{=:B(\mathsf{L})^{-1}} s_{i,t} \\ &= \left(1 - \frac{\mathsf{L}(\rho - \lambda)}{1 - \lambda\mathsf{L}}\right) s_{i,t} \\ &= \frac{1 - \rho\mathsf{L}}{1 - \lambda\mathsf{L}} s_{i,t} \end{aligned}$$

Hence, we have:

$$B(\mathsf{L}) = \frac{1 - \lambda\mathsf{L}}{1 - \rho\mathsf{L}}.$$

The forecast error is distributed as:

$$w_{i,t} \sim \mathcal{N}\big(0,\, V\big), \qquad V \equiv HPH^\top + \sigma_u^2 = \sigma_u^2 \frac{\rho}{\lambda}$$

A.4 Proof of proposition 1

The proof of the proposition builds on the framework developed by Huo and Takayama (2025) for solving rational expectation equilibria under incomplete information.

Wiener-Hopf and Annihilation. We begin by reminding the reader of the Wiener-Hopf prediction formula for the general case. Let $f_{i,t} \equiv A(\mathsf{L})\eta_{i,t}$. The Wiener-Hopf formula is:

$$\begin{aligned} \mathbb{E}_{i,t}(f_{i,t}) &= \lfloor A(\mathsf{L})M^\top(\mathsf{L}^{-1})B^\top(\mathsf{L}^{-1})^{-1} \rfloor_+ V^{-1} w_{i,t} \\ &= \lfloor A(\mathsf{L})M^\top(\mathsf{L}^{-1})B^\top(\mathsf{L}^{-1})^{-1} \rfloor_+ V^{-1} B(\mathsf{L})^{-1} s_{i,t} \end{aligned}$$

For any function $g(\mathsf{L})$ that has no negative roots in expansion, and when the roots $|\vartheta_k| < 1$ are not repeated, the annihilator operator is calculated as:

$$\left\lfloor \frac{g(\mathsf{L})}{\prod_{k=1}^{\ell}(\mathsf{L} - \vartheta_k)} \right\rfloor_+ = \frac{g(\mathsf{L})}{\prod_{k=1}^{\ell}(\mathsf{L} - \vartheta_k)} - \sum_{k=1}^{\ell} \left(\frac{1}{\mathsf{L} - \vartheta_k}\right) \frac{g(z)}{\prod_{h \neq k}(z - \vartheta_h)} \bigg|_{z=\vartheta_k}$$

For any function $g(\mathsf{L})$ that has no negative roots in expansion, and when the roots $|\vartheta_k| < 1$ are repeated α_k times, the annihilator operator is calculated as:

$$\left\lfloor \frac{g(\mathsf{L})}{\prod_{k=1}^{\ell}(\mathsf{L} - \vartheta_k)^{\alpha_k}} \right\rfloor_+ = \frac{g(\mathsf{L})}{\prod_{k=1}^{\ell}(\mathsf{L} - \vartheta_k)^{\alpha_k}} - \sum_{k=1}^{\ell} \sum_{\tau=0}^{\alpha_k - 1} \frac{1}{\tau!} \left(\frac{1}{\mathsf{L} - \vartheta_k}\right)^{\alpha_k - \tau} \left(\frac{d^\tau}{dz^\tau} \frac{g(z)}{\prod_{h \neq k}(z - \vartheta_h)^{\alpha_h}} \bigg|_{z=\vartheta_k} \right)$$

Guess of the solution. We guess a solution of the following form:

$$\mathbb{E}_{i,t}(\Delta p_{i,t} - \pi_t) = h(\mathsf{L})\widehat{s_{i,t}}$$

for some causal (analytical) lag polynomial $h(\mathsf{L})$, where the rescaled signal is defined as $\widehat{s_{i,t}} \equiv \sigma_\xi^2/\sigma_\varepsilon^2\, s_{i,t}$. The solution is obtained by computing the right-hand side of equation (12) term by term and "matching the coefficients" with those of the guess.

Contemporaneous term. We seek now to compute:

$$\mathbb{E}_{i,t}(p_{i,t} - p_t) = \mathbb{E}_{i,t}\left(\frac{h(\mathsf{L})}{(1-\mathsf{L})}\widehat{s_{i,t}}\right) = \mathbb{E}_{i,t}\left(\frac{h(\mathsf{L})}{(1-\mathsf{L})}M(\mathsf{L})\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\eta_{i,t}\right) = \frac{h(\mathsf{L})}{(1-\mathsf{L})}\widehat{s_{i,t}}$$

because $h(\mathsf{L})$, $1/(1-\mathsf{L})$, and $M(\mathsf{L})$ are causal polynomials.

Future price. As the future price involves a non-causal lag polynomial, we use the Wiener-Hopf projection formula:

$$\begin{aligned}
\mathbb{E}_{i,t}(\Delta p_{i,t+1} - \pi_{t+1}) &= \mathbb{E}_{i,t}\left(\frac{h(\mathsf{L})}{\mathsf{L}}M(\mathsf{L})\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\eta_{i,t}\right) \\
&= \left\lfloor \frac{h(\mathsf{L})}{\mathsf{L}} M(\mathsf{L})M^\top(\mathsf{L}^{-1})(B^\top(\mathsf{L}^{-1}))^{-1} \right\rfloor_+ V^{-1}B(\mathsf{L})^{-1}\widehat{s_{i,t}} \\
&= \left\lfloor \frac{h(\mathsf{L})}{\mathsf{L}} B(\mathsf{L})VB^\top(\mathsf{L}^{-1})(B^\top(\mathsf{L}^{-1}))^{-1} \right\rfloor_+ V^{-1}B(\mathsf{L})^{-1}\widehat{s_{i,t}} \\
&= \left\lfloor h(\mathsf{L})B(\mathsf{L})\,\frac{1}{\mathsf{L}} \right\rfloor_+ B(\mathsf{L})^{-1}\widehat{s_{i,t}} \\
&= \Bigg\lfloor \underbrace{\frac{h(\mathsf{L})(1-\lambda\mathsf{L})}{(1-\rho\mathsf{L})}}_{\text{causal}} \underbrace{\frac{1}{\mathsf{L}}}_{\text{not causal}} \Bigg\rfloor_+ B(\mathsf{L})^{-1}\widehat{s_{i,t}} \\
&= \left(h(\mathsf{L})\mathsf{L}^{-1} - \frac{h(0)(1-\rho\mathsf{L})}{\mathsf{L}(1-\lambda\mathsf{L})}\right)\widehat{s_{i,t}}
\end{aligned}$$

Verification. Plugging into the best response:

$$\mathbb{E}_{i,t}(\Delta p_{i,t} - \pi_t) = \mathbb{E}_{i,t}\left\{\kappa\Big((1-\Omega)\xi_{i,t} - (p_{i,t} - p_t)\Big) + \beta(\Delta p_{i,t+1} - \pi_{t+1})\right\}$$

and canceling $\widehat{s_{i,t}}$ leads to:

$$h(\mathsf{L}) = \kappa\left[(1-\Omega)\frac{1}{1-\lambda\mathsf{L}}\left(1-\frac{\lambda}{\rho}\right) - \frac{h(\mathsf{L})}{(1-\mathsf{L})}\right] + \beta\left(h(\mathsf{L})\mathsf{L}^{-1} - \frac{h(0)(1-\rho\mathsf{L})}{\mathsf{L}(1-\lambda\mathsf{L})}\right)$$

Grouping terms and multiplying by the common denominator $(1-\mathsf{L})\mathsf{L}(1-\lambda\mathsf{L})$:

$$\mathcal{C}(z)\cdot h(z) = \mathcal{D}(z; h(0))$$

where

$$\mathcal{C}(z) \equiv (1-\lambda z)\Big((1-z)z + \kappa z - \beta(1-z)\Big)$$

and

$$\mathcal{D}(z; h(0)) \equiv \ \kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)(1-z)z - \beta h(0)(1-\rho z)(1-z)$$

The roots of $\mathcal{C}(z)$ are

$$z \in \Big\{ \underbrace{\frac{1}{\lambda},\ \vartheta,}_{\text{Outside}}\ \underbrace{\frac{\beta}{\vartheta}}_{\text{Inside}} \Big\}.$$

where

$$\vartheta = \frac{(1+\beta+\kappa) + \sqrt{(1+\beta+\kappa)^2 - 4\beta}}{2} > 1$$

The roots of $\mathcal{D}(z, h(0))$ are:

$$z_1 = 1, \qquad z_2 = \frac{\beta h(0)}{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right) + \beta h(0)\rho}.$$

We pick the constant $h(0)$ so that it exactly cancels with the inside root β/ϑ. This step ensures that the solution $h(z)$ is a causal lag polynomial. In turn,

$$h(z) = \frac{\mathcal{D}(z)}{\mathcal{C}(z)} = \frac{(1-z)K}{(z-\vartheta)(z-1/\lambda)},$$

where

$$K \equiv \frac{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)\vartheta}{(\vartheta-\rho\beta)\lambda}.$$

Hence $h(z)$ is an ARMA(2,1) process and $\mathbb{E}_{i,t}(p_{i,t}-p_t)$ is an AR(2) process in $\widehat{s}_{i,t}$:

$$\mathbb{E}_{i,t}(p_{i,t}-p_t) = \frac{K}{(\mathsf{L}-\vartheta)(\mathsf{L}-1/\lambda)}\widehat{s}_{i,t}$$

A.5 Proof of proposition 2

Consider the following reduced-form pricing equation:

$$\mathbb{E}_t(\Delta\widehat{p}_{i,t}) = \kappa\Big[(1-\Omega)\xi_{i,t} - \omega_p\ \mathbb{E}_t(\widehat{p}_{i,t})\Big] + \omega_f\ \mathbb{E}_t\left\{\Delta\widehat{p}_{i,t+1}\right\} + \omega_\ell\Delta\widehat{p}_{i,t-1}$$

where the hat denotes deviations from the average, e.g. $\widehat{p}_{i,t} \equiv p_{i,t} - p_t$ and $\xi_{i,t}$, $\Delta\widehat{p}_{i,t-1} \in \mathcal{I}_t$ by definition of complete information. The expectation is taken with respect to the distribution of the i.i.d. Bernoulli price adjustment opportunity. We guess that the price satisfies the following ARMA(2,1) process:

$$\widehat{p}_{i,t} = A\xi_{i,t} + B\widehat{p}_{i,t-1} + C\widehat{p}_{i,t-2} + \psi_{i,t} \tag{A.5}$$

where $\psi_{i,t} \overset{\text{i.i.d.}}{\sim} \mathcal{B}(1-\theta)$ is a mean-zero Bernoulli error term with intensity $(1-\theta)$. Without loss of generality, we ignore the error term. Using the guess:

$$\Delta\widehat{p}_{i,t} = A\xi_{i,t} + (B-1)\widehat{p}_{i,t-1} + C\widehat{p}_{i,t-2}$$

$$\begin{aligned}\mathbb{E}_t(\Delta\widehat{p}_{i,t+1}) &= A\rho\xi_{i,t} + (B-1)\widehat{p}_{i,t} + C\widehat{p}_{i,t-1} \\ &= A(\rho + (B-1))\xi_{i,t} + ((B-1)B + C)\widehat{p}_{i,t-1} + (B-1)C\widehat{p}_{i,t-2}\end{aligned}$$

Replace into the best response and match the coefficients:

$$A = \kappa\Big((1-\Omega) - \omega_p A\Big) + \omega_f A(\rho + (B-1))$$

$$B = 1 - \kappa\omega_p B + \omega_f\Big((B-1)B + C\Big) + \omega_\ell$$

$$C = -\kappa\omega_p C + \omega_f(B-1)C - \omega_\ell$$

The system can be rewritten as:

$$\begin{aligned}A &= \frac{\kappa(1-\Omega)}{1 + \kappa\omega_p - \omega_f(\rho + B - 1)} \\ C &= \frac{-\omega_\ell}{1 + \kappa\omega_p - \omega_f(B-1)} \\ 0 &= \omega_f^2 B^3 - 2\,\omega_f\Big(\omega_f + \kappa\omega_p + 1\Big)B^2 \\ &\quad + \Big((\omega_f)^2 + 2\kappa\omega_f\omega_p + \omega_f\,\omega_\ell + 3\omega_f + \kappa^2\omega_p^2 + 2\kappa\omega_p + 1\Big)B \\ &\quad - \Big(\omega_f + \kappa\omega_\ell\omega_p + \kappa\omega_p + \omega_\ell + 1\Big).\end{aligned} \tag{A.6}$$

Denote by $B_\star$ a real solution to the cubic equation. We can rewrite the solution as:

$$(\mathsf{L} - \mathcal{R}_1)(\mathsf{L} - \mathcal{R}_2)\widehat{p}_{i,t} = -\frac{A}{C}\,\xi_{i,t}$$

with

$$C = -\frac{1}{\mathcal{R}_1 \mathcal{R}_2}, \qquad B_\star = -C(\mathcal{R}_1 + \mathcal{R}_2).$$

Using $\mathbb{E}(s_{i,t} \mid \{\xi_{i,\tau}\}_{\tau \leq t}) = \mathbb{E}(\xi_{i,t})$ because the unconditional expectation $\mathbb{E}(mc_t) = 0$, the observational equivalence (condition 16) requires that:

$$h(z)\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\xi_{i,t} = h^{\mathrm{CI}}(z)\xi_{i,t}$$

where $h(z)$ is the structural lag polynomial from proposition 1 and $h^{\mathrm{CI}}(\mathsf{L})$ is the ARMA(2,1) process in equation (A.5). The condition is satisfied if and only if:

$$-\frac{A}{C} = \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}K, \qquad \mathcal{R}_1 = \vartheta, \qquad \mathcal{R}_2 = \frac{1}{\lambda}$$

Hence, it holds:

$$A = \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\frac{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)}{(\vartheta - \rho\beta)}, \qquad C = -\frac{\lambda}{\vartheta}, \qquad B_\star = \lambda + \frac{1}{\vartheta}.$$

We replace the above conditions into the system of equations (A.6) and solve for the parameters ω_p, ω_f, and ω_ℓ. This leads to the formulas in the text.

A.6 Proof of corollary 1

Define the signal–to–noise ratio $r \equiv \sigma_\varepsilon^2/\sigma_u^2$. The steady-state Riccati equation implies

$$\lambda^2 - \left(\rho + \frac{1+r}{\rho}\right)\lambda + 1 = 0 \quad \Longleftrightarrow \quad \lambda(r) = \frac{A(r) - \sqrt{A(r)^2 - 4}}{2}, \;\; A(r) = \rho + \frac{1+r}{\rho},$$

and equivalently

$$r = \rho\left(\lambda + \frac{1}{\lambda}\right) - \rho^2 - 1.$$

Since $A'(r) = 1/\rho > 0$ and $A \mapsto (A - \sqrt{A^2-4})/2$ is strictly decreasing on $(2, \infty)$, $\lambda(r)$ is strictly decreasing in r with $\lim_{r\to\infty}\lambda(r) = 0$. Let

$$D \equiv (\rho - \lambda)(\lambda\rho - \lambda - \rho - \lambda\rho\vartheta).$$

Because $\rho > \lambda > 0$ and $\vartheta > 1$, one has $D < 0$.

(i) Positivity of ω_ℓ. From the closed form,

$$\omega_\ell = \frac{\lambda\rho^2(\beta\lambda - \vartheta)}{\vartheta(\lambda-\rho)(\lambda\rho\vartheta - \lambda\rho + \lambda + \rho)}.$$

Note that $\beta\lambda - \vartheta < 0$, $\lambda - \rho < 0$, and $\lambda\rho\vartheta - \lambda\rho + \lambda + \rho > 0$. Hence the numerator and denominator are both negative, which implies $\omega_\ell > 0$. Moreover, as $r \to \infty$ (so $\lambda \to 0$),

$$\omega_\ell \longrightarrow 0.$$

(ii) Upper bound for ω_f. Dividing by ω_f by β:

$$\frac{\omega_f}{\beta} = \frac{1}{\beta\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)D}\left\{\lambda\vartheta\Big[\rho(\vartheta-1)+1\Big] + \beta\rho^2(\lambda - 1 - \lambda\vartheta) + \left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} - 1\right)\vartheta(\lambda-\rho)\right\}.$$

Taking the complete-information limit $r \to \infty$ (so $\lambda \to 0$) yields

$$\lim_{r\to\infty}\frac{\omega_f}{\beta} = \frac{\beta\rho + \vartheta\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} - 1\right)}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\beta\rho} = 1 - \frac{\left(1 - \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)(\vartheta - \beta\rho)}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\beta\rho} \;\le\; 1,$$

because $\sigma_\xi^2/\sigma_\varepsilon^2 \le 1$ and $\vartheta - \beta\rho > 0$ (since $\vartheta > 1$ while $\beta\rho < 1$). Multiplying by β, we obtain

$$\lim_{r\to\infty}\omega_f \;\le\; \beta, \qquad \lim_{\substack{r\to\infty \\ \sigma_\xi^2/\sigma_\varepsilon^2 \to 1}} \omega_f \;=\; \beta.$$

Next, the gap between the complete-information limit and the ω_f satisfies:

$$\lim_{r\to\infty}\frac{\omega_f}{\beta} - \frac{\omega_f}{\beta} = \frac{\lambda\left(1 + \rho(\vartheta-1)\right)\left[\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\vartheta(\rho-\lambda) + \lambda(\vartheta - \beta\rho)\right]}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\beta\rho\,(\rho-\lambda)(\lambda\rho\vartheta - \lambda\rho + \lambda + \rho)}.$$

All factors in the numerator and denominator are positive for any $\lambda \in (0, \rho)$, because $\rho - \lambda > 0$, $\vartheta - \beta\rho > 0$, and $\lambda\rho\vartheta - \lambda\rho + \lambda + \rho = \lambda\rho(\vartheta - 1) + \lambda + \rho > 0$. Therefore

$$\omega_f \;<\; \lim_{r\to\infty}\omega_f \qquad \text{for any finite } r.$$

(iii) Lower bound for ω_p. From the formula for ω_p,

$$\omega_p = \frac{(1-\lambda)(1-\vartheta)}{\kappa\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)D}\left\{\lambda\rho(\vartheta-1)+\lambda+\rho^2-\beta\rho^2\left(\lambda+\frac{1}{\vartheta}\right)+\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}-1\right)(\lambda+\rho^2-\rho-\lambda\rho)\right\}.$$

Imposing the identity $\kappa = \vartheta - 1 - \beta + \beta/\vartheta$, the complete-information limit $r \to \infty$ (so $\lambda \to 0$) satisfies

$$\lim_{r\to\infty}\omega_p - 1 = \frac{\left(1-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)(\vartheta-\beta\rho)}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\rho\,(\vartheta-\beta)} \;\geq\; 0,$$

because $\sigma_\xi^2/\sigma_\varepsilon^2 \leq 1$, $\vartheta - \beta\rho > 0$, and $\vartheta - \beta > 0$. Hence

$$\lim_{r\to\infty}\omega_p \;\geq\; 1, \qquad \lim_{\substack{r\to\infty \\ \sigma_\xi^2/\sigma_\varepsilon^2\to 1}} \omega_p \;=\; 1.$$

Moreover, the deviation from this complete-information limit satisfies:

$$\omega_p - \lim_{r\to\infty}\omega_p = \frac{\lambda\,(1-\rho)\,(\vartheta-1)\,(\rho\vartheta+1)\left[\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\vartheta(\rho-\lambda)+\lambda(\vartheta-\beta\rho)\right]}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\right)\kappa\rho\vartheta\,(\rho-\lambda)\,(\lambda\rho\vartheta-\lambda\rho+\lambda+\rho)}.$$

All factors are positive for any $\lambda \in (0,\rho)$ and $\rho \in (0,1)$, implying

$$\omega_p \;>\; \lim_{r\to\infty}\omega_p \qquad \text{for any finite } r.$$

A.7 Proof of proposition 3

The pricing equation for the aggregate components is:

$$\mathbb{E}_{i,t}\,\pi_t = \mathbb{E}_{i,t}\{\kappa(1-\Omega)mc_t + \beta\pi_{t+1}\}$$

Guess and verify a solution of the form:

$$\mathbb{E}_{i,t}(\pi_t) = h^a(\mathsf{L})\frac{\sigma_a^2}{\sigma_\varepsilon^2}s_{i,t}$$

Plugging into the pricing equation and using the Wiener-Hopf prediction formula (same steps as the proof of proposition 1), we get:

$$h^a(\mathsf{L}) = \kappa\left[(1-\Omega)\frac{1}{1-\lambda\mathsf{L}}\left(1-\frac{\lambda}{\rho}\right)\right] + \beta\left(h^a(\mathsf{L})\mathsf{L}^{-1} - \frac{h^a(0)(1-\rho\mathsf{L})}{\mathsf{L}(1-\lambda\mathsf{L})}\right)$$

Grouping terms and multiplying by the common denominator $\mathsf{L}(1-\lambda\mathsf{L})$:

$$\mathcal{C}(z)\cdot h^a(z) = \mathcal{D}(z; h^a(0))$$

where

$$\mathcal{C}(z) \equiv (1-\lambda z)(z-\beta)$$

and

$$\mathcal{D}(z; h^a(0)) \equiv \ \kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)z - \beta h^a(0)(1-\rho z).$$

The roots of $\mathcal{C}(z)$ are

$$z \in \Big\{ \underbrace{\frac{1}{\lambda}}_{\text{Outside}}, \underbrace{\beta}_{\text{Inside}} \Big\}.$$

The roots of $\mathcal{D}(z, h^a(0))$ are:

$$z_1 = \frac{\beta h^a(0)}{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right) + \beta h^a(0)\rho}.$$

We pick the constant $h^a(0)$ so that it exactly cancels with the inside root β. In turn,

$$h^a(z) = \frac{\mathcal{D}(z)}{\mathcal{C}(z)} = \frac{K^a}{z - 1/\lambda},$$

where

$$K^a \equiv -\frac{\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)}{\lambda(1-\rho\beta)}.$$

Hence $h^a(z)$ is an AR(1) process and $\mathbb{E}_{i,t}\,\pi_t$ is an AR(1) process in $\frac{\sigma_a^2}{\sigma_\varepsilon^2}s_{i,t}$:

$$\mathbb{E}_{i,t}\,\pi_t = \frac{K^a}{(\mathsf{L}-1/\lambda)}\frac{\sigma_a^2}{\sigma_\varepsilon^2}s_{i,t}$$

Summing the two components, we get:

$$\begin{aligned}
\mathbb{E}_{i,t}(\Delta p_{i,t}) &= \left(\frac{K^a}{(\mathsf{L}-1/\lambda)} \frac{\sigma_a^2}{\sigma_\varepsilon^2} + \frac{(1-\mathsf{L})K}{(\mathsf{L}-\vartheta)(\mathsf{L}-1/\lambda)} \frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) s_{i,t} \\
&= \frac{\left[\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \vartheta \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a \right] - \left[\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a \right] \mathsf{L}}{(\mathsf{L}-\vartheta)(\mathsf{L}-1/\lambda)} s_{i,t} \\
&= \frac{\left[\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \vartheta \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a \right] (1-\gamma \mathsf{L})}{(\mathsf{L}-\vartheta)(\mathsf{L}-1/\lambda)} s_{i,t}, \\
&= \frac{K^s (1-\gamma \mathsf{L})}{(\mathsf{L}-\vartheta)(\mathsf{L}-1/\lambda)} s_{i,t},
\end{aligned}$$

where the MA root is:

$$\gamma \equiv \frac{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a}{\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \vartheta \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a}$$

and

$$K^s \equiv \left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} \right) K - \vartheta \left(\frac{\sigma_a^2}{\sigma_\varepsilon^2} \right) K^a = \frac{\kappa(1-\Omega)(1-\lambda/\rho)\vartheta}{\lambda(\gamma\vartheta - \rho\beta)}.$$

Hence, $\mathbb{E}_{i,t}(\Delta p_{i,t})$ follows an ARMA(2,1) process with AR roots $\vartheta, 1/\lambda$ and MA root γ. Simplifying the formulas above and rearranging leads to the equation (23) in the text.

A.8 Proof of proposition 4

Consider the following "distorted" pricing equation:

$$\Delta p_{i,t} = \kappa(1-\Omega) mc_{i,t} + \omega_c^a mc_{i,t-1} + \omega_f^a \mathbb{E}_t \{\Delta p_{i,t+1}\} + \omega_{\ell,1}^a \Delta p_{i,t-1} + \omega_{\ell,2}^a \Delta p_{i,t-2}. \tag{A.7}$$

We guess a solution of the form:

$$\Delta p_{i,t} = A mc_{i,t} + B mc_{i,t-1} + C \Delta p_{i,t-1} + D \Delta p_{i,t-2}$$

Shifting the guess forward by one period,

$$\Delta p_{i,t+1} = A mc_{i,t+1} + B mc_{i,t} + C \Delta p_{i,t} + D \Delta p_{i,t-1},$$

and taking expectations yields

$$\mathbb{E}_t(\Delta p_{i,t+1}) = A\rho\, mc_{i,t} + B mc_{i,t} + C\Delta p_{i,t} + D\Delta p_{i,t-1}.$$

Using the conjecture to substitute $\Delta p_{i,t}$ inside the expectation, we obtain

$$\mathbb{E}_t(\Delta p_{i,t+1}) = \Big[A(\rho + C) + B\Big] mc_{i,t} + CB\, mc_{i,t-1} + (C^2 + D)\Delta p_{i,t-1} + CD\,\Delta p_{i,t-2}.$$

Plugging into the distorted pricing equation and matching coefficients gives the system:

$$\begin{aligned} A &= \kappa(1-\Omega) + \omega_f^a\Big(A(\rho + C) + B\Big), \\ B &= \omega_c^a + \omega_f^a\, CB, \\ C &= \omega_{\ell,1}^a + \omega_f^a\,(C^2 + D), \\ D &= \omega_{\ell,2}^a + \omega_f^a\, CD. \end{aligned} \tag{A.8}$$

Observational equivalence. The distorted pricing equation takes the form:

$$\Big(1 - C\mathsf{L} - D\mathsf{L}^2\Big)\Delta p_{i,t} \;=\; (A + B\mathsf{L})\, mc_{i,t}.$$

Using the observational equivalence (equation 24) and matching the coefficients with the model solution leads to:

$$1 - Cz - Dz^2 \;=\; (1 - z/\vartheta)(1 - \lambda z), \qquad (A + Bz) \;=\; \frac{\kappa(1-\Omega)(1-\lambda/\rho)}{(\gamma\vartheta - \rho\beta)}(1 - \gamma z).$$

Therefore we get the system of restrictions:

$$C \;=\; \lambda + \frac{1}{\vartheta}, \qquad D \;=\; -\frac{\lambda}{\vartheta}, \qquad A \;=\; \frac{\kappa(1-\Omega)(1-\lambda/\rho)}{(\gamma\vartheta - \rho\beta)}, \qquad B \;=\; -\frac{\kappa(1-\Omega)(1-\lambda/\rho)}{(\gamma\vartheta - \rho\beta)}\,\gamma,$$

Plugging these restrictions into (A.8) and rearranging, we get to the formulas for the coefficients in the text. Finally, the Phillips curve is obtained by integrating the pricing equation (A.7) in the cross-section.

Sign restrictions for the coefficients. First, we show $\omega_f^a < \beta$. Using the definition of ω_f^a,

$$\beta - \omega_f^a = \frac{\beta(\rho - \lambda)\Big(\rho + \lambda + \frac{1}{\vartheta} - \gamma\Big) - \Big(\rho + \beta\rho^2 - \gamma\rho\vartheta - \lambda\Big)}{(\rho - \lambda)\Big(\rho + \lambda + \frac{1}{\vartheta} - \gamma\Big)}.$$

Rearranging the numerator gives

$$\beta(\rho-\lambda)\left(\rho+\lambda+\frac{1}{\vartheta}-\gamma\right)-\left(\rho+\beta\rho^2-\gamma\rho\vartheta-\lambda\right)=\lambda(1-\beta\lambda)+\left(\gamma-\frac{1}{\vartheta}\right)\left(\rho(\vartheta-\beta)+\beta\lambda\right).$$

Each term on the right-hand side is nonnegative, and at least one is strictly positive: $\lambda(1-\beta\lambda)>0$ because $\lambda\in(0,1)$, $\beta\in(0,1)$, and $\gamma-\frac{1}{\vartheta}\geq 0$. Since the denominator is positive, we obtain $\beta-\omega_f^a>0$.

Moreover, we show $1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)>0$. Using the definition of ω_f^a again,

$$1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)=\frac{(\rho-\lambda)\left(\rho+\lambda+\frac{1}{\vartheta}-\gamma\right)-\left(\lambda+\frac{1}{\vartheta}\right)\left(\rho+\beta\rho^2-\gamma\rho\vartheta-\lambda\right)}{(\rho-\lambda)\left(\rho+\lambda+\frac{1}{\vartheta}-\gamma\right)}.$$

Rearranging the numerator yields

$$(\rho-\lambda)\left(\rho+\lambda+\frac{1}{\vartheta}-\gamma\right)-\left(\lambda+\frac{1}{\vartheta}\right)\left(\rho+\beta\rho^2-\gamma\rho\vartheta-\lambda\right)=\frac{\rho^2(\vartheta-\beta)+\lambda(1-\beta\rho^2\vartheta)}{\vartheta}+\left(\gamma-\frac{1}{\vartheta}\right)\lambda(\rho\vartheta+1).$$

The second term is nonnegative because $\gamma\geq\frac{1}{\vartheta}$. For the first term, note that it is affine in λ on the interval $\lambda\in(0,\rho)$. Hence its minimum over $\lambda\in[0,\rho]$ is attained at one of the endpoints:

$$\rho^2(\vartheta-\beta)+\lambda(1-\beta\rho^2\vartheta)\ \geq\ \min\left\{\rho^2(\vartheta-\beta),\ \rho^2(\vartheta-\beta)+\rho(1-\beta\rho^2\vartheta)\right\}.$$

The first endpoint is strictly positive because $\vartheta>\beta$. The second endpoint is rearranged as

$$\rho^2(\vartheta-\beta)+\rho(1-\beta\rho^2\vartheta)=-\rho(\beta\rho-1)(\rho\vartheta+1),$$

which is strictly positive because $\beta\rho<1$ and $\rho\vartheta+1>0$. Therefore, the entire numerator is strictly positive, and the denominator is positive.

Second, we show $\omega_c^a<0$. By definition,

$$\omega_c^a=-\kappa(1-\Omega)\left(1-\frac{\lambda}{\rho}\right)\frac{\gamma}{\gamma\vartheta-\rho\beta}\left(1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)\right).$$

All factors multiplying are strictly positive from the previous step.

Third, we show $\omega_{\ell,2}^a<0$. This follows from the definition:

$$\omega_{\ell,2}^a=-\left(\frac{\lambda}{\vartheta}\right)\left(1-\omega_f^a\left(\lambda+\frac{1}{\vartheta}\right)\right)<0$$

because $\lambda/\vartheta>0$ and $1-\omega_f^a(\lambda+1/\vartheta)>0$.

Fourth, we show $\omega^a_{\ell,1} > 0$. Starting from the definition of $\omega^a_{\ell,1}$ and using algebra,

$$\begin{aligned}\omega^a_{\ell,1} &= \left(\lambda + \frac{1}{\vartheta}\right)\left(1 - \omega^a_f\left(\lambda + \frac{1}{\vartheta}\right)\right) + \omega^a_f\left(\frac{\lambda}{\vartheta}\right) \\ &= \frac{\left(1 - \omega^a_f\left(\lambda + \frac{1}{\vartheta}\right)\right)\left[\left(\lambda + \frac{1}{\vartheta}\right)^2 - \frac{\lambda}{\vartheta}\right] + \frac{\lambda}{\vartheta}}{\lambda + \frac{1}{\vartheta}}.\end{aligned}$$

The denominator is positive. In the numerator,

$$\left(\lambda + \frac{1}{\vartheta}\right)^2 - \frac{\lambda}{\vartheta} = \lambda^2 + \frac{\lambda}{\vartheta} + \frac{1}{\vartheta^2} > 0,$$

and $1 - \omega^a_f(\lambda + 1/\vartheta) > 0$ as shown above, so the numerator is strictly larger than $\lambda/\vartheta > 0$.

A.9 Proof of Corollary 3

Let $x \equiv \sigma^2_\xi / \sigma^2_\varepsilon \in (0,1)$ and define

$$\Delta(x) \;\equiv\; \omega^a_f(x) - \omega_f(x).$$

Monotonicity of γ. From proposition 3,

$$\gamma(x) = \frac{x(1-\rho\beta) + (1-x)\frac{\vartheta - \rho\beta}{\vartheta}}{x(1-\rho\beta) + (1-x)(\vartheta - \rho\beta)}.$$

Taking derivatives with respect to x:

$$\gamma'(x) = \frac{(1-\rho\beta)(\vartheta - \rho\beta)(\vartheta - 1)}{\vartheta(x(1-\rho\beta) + (1-x)(\vartheta - \rho\beta))^2},$$

where

$$(1-\rho\beta)(\vartheta - \rho\beta)(\vartheta - 1) > 0,$$

so $\gamma'(x) > 0$. Hence $\gamma(x)$ is strictly increasing in x.

Monotonicity of ω^a_f From proposition 4,

$$\omega^a_f = \frac{\rho + \beta\rho^2 - \lambda - \rho\vartheta\gamma}{(\rho - \lambda)(\rho + \lambda + 1/\vartheta - \gamma)}.$$

Let $H(\gamma) \equiv \rho + \lambda + 1/\vartheta - \gamma > 0$ and $N(\gamma) \equiv \rho + \beta\rho^2 - \lambda - \rho\vartheta\gamma$. Then $\omega_f^a(\gamma) = \frac{N(\gamma)}{(\rho-\lambda)H(\gamma)}$ and

$$\frac{\partial \omega_f^a}{\partial \gamma} = \frac{1}{\rho - \lambda} \cdot \frac{N'(\gamma)H(\gamma) - N(\gamma)H'(\gamma)}{H(\gamma)^2}.$$

Using $N'(\gamma) = -\rho\vartheta$ and $H'(\gamma) = -1$, this becomes

$$\frac{\partial \omega_f^a}{\partial \gamma} = \frac{-\rho\vartheta H(\gamma) + N(\gamma)}{(\rho - \lambda)H(\gamma)^2}.$$

Substituting $H(\gamma) = \rho + \lambda + 1/\vartheta - \gamma$ and $N(\gamma) = \rho + \beta\rho^2 - \lambda - \rho\vartheta\gamma$,

$$\begin{aligned} -\rho\vartheta H(\gamma) + N(\gamma) &= -\rho\vartheta(\rho + \lambda + 1/\vartheta - \gamma) + \rho + \beta\rho^2 - \lambda - \rho\vartheta\gamma \\ &= -\rho\vartheta(\rho + \lambda + 1/\vartheta) + \rho + \beta\rho^2 - \lambda \\ &= \beta\rho^2 - \rho^2\vartheta - \lambda\rho\vartheta - \lambda. \end{aligned}$$

Since $\vartheta > 1$ and $\lambda > 0$,

$$\beta\rho^2 - \rho^2\vartheta - \lambda\rho\vartheta - \lambda = \rho^2(\beta - \vartheta) - \lambda(\rho\vartheta + 1) < 0.$$

Also $\rho - \lambda > 0$ and $H(\gamma)^2 > 0$, so $\partial\omega_f^a/\partial\gamma < 0$. Combined with the monotonicity of γ ($\gamma'(x) > 0$), we conclude that $\omega_f^a(x)$ is strictly decreasing in x.

Monotonicity of ω_f. From proposition 2,

$$\omega_f(x) = \frac{\lambda\vartheta[\rho(\vartheta - 1) + 1] + \beta\rho^2(\lambda - 1 - \lambda\vartheta) + (x-1)\vartheta(\lambda - \rho)}{xD}, \qquad D \equiv (\rho - \lambda)(\lambda\rho - \lambda - \rho - \lambda\vartheta).$$

Let

$$A \equiv \lambda\vartheta[\rho(\vartheta - 1) + 1] + \beta\rho^2(\lambda - 1 - \lambda\vartheta), \qquad B \equiv \vartheta(\lambda - \rho),$$

so $\omega_f(x) = \frac{A + (x-1)B}{xD}$. Differentiating,

$$\omega_f'(x) = \frac{B \cdot (\beta x D) - (A + (x-1)B) \cdot (\beta D)}{(xD)^2} = \frac{B - A}{x^2 D}.$$

A simplification yields

$$B - A = \rho(\beta\rho - \vartheta)\Big(1 + \lambda(\vartheta - 1)\Big) < 0,$$

because $\beta\rho < 1 < \vartheta$ implies $\beta\rho - \vartheta < 0$ and $1 + \lambda(\vartheta - 1) > 0$. Moreover,

$$D = (\rho - \lambda)\Big(\lambda\rho - \lambda - \rho - \lambda\vartheta\Big) = -(\rho - \lambda)\Big(\rho + \lambda(\vartheta + 1 - \rho)\Big) < 0,$$

since $\rho - \lambda > 0$ and $\rho + \lambda(\vartheta + 1 - \rho) > 0$. Therefore $\omega_f'(x) > 0$: $\omega_f(x)$ is strictly increasing in x.

Threshold. By the previous steps, $\Delta(x) = \omega_f^a(x) - \omega_f(x)$ is strictly decreasing and continuous on $(0,1)$. Also, as $x \downarrow 0$,

$$\omega_f(x) \to \frac{A-B}{xD}, \qquad A - B > 0,$$

while $xD \to 0^-$ because $D < 0$. Hence $\omega_f(x) \to -\infty$, and since $\omega_f^a(x)$ is bounded, $\Delta(x) \to +\infty$ as $x \downarrow 0$. Define

$$r^\star \equiv \inf\{x \in (0,1) : \Delta(x) < 0\} \in (0,1].$$

Because Δ is strictly decreasing, we have $\Delta(x) < 0$ if and only if $x > r^\star$, i.e.

$$\omega_f^a < \omega_f \quad \Longleftrightarrow \quad \frac{\sigma_\xi^2}{\sigma_\varepsilon^2} > r^\star.$$

A.10 Implementation of the baseline specification

We estimate the parameters θ, Ω, ω_f, and ω_p of the moment condition (19) and impose the parameter restrictions on ω_ℓ, as follows. First, we recover ρ from:

$$a_\rho \rho^2 + b_\rho \rho + c_\rho = 0$$

where

$$\begin{aligned}
a_\rho &= \vartheta(\vartheta - 1)\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_f - \beta\right), \\
b_\rho &= (\vartheta - 1)\left[(\vartheta - 1)\left(\vartheta - \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_f\right) - \kappa\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_p\,\vartheta - \vartheta\left(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2} - 1\right)\right], \\
c_\rho &= \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\left[(\vartheta - 1)(\vartheta - \omega_f) - \kappa\omega_p\,\vartheta\right]
\end{aligned}$$

The admissible root is:

$$\rho = \frac{-b_\rho - \sqrt{b_\rho^2 - 4a_\rho c_\rho}}{2a_\rho}.$$

Second, we recover λ from:

$$a_\lambda \lambda^2 + b_\lambda \lambda + c_\lambda = 0,$$

where

$$
\begin{aligned}
a_\lambda &= \frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_f\Big[1+\rho(\vartheta-1)\Big],\\
b_\lambda &= \Big(\beta-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_f\Big)\rho^2(\vartheta-1)-\vartheta\Big[\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}+\rho(\vartheta-1)\Big],\\
c_\lambda &= \Big(\beta-\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}\omega_f\Big)\rho^2+\vartheta\Big(\frac{\sigma_\xi^2}{\sigma_\varepsilon^2}-1\Big)\rho.
\end{aligned}
$$

The admissible root is:

$$
\lambda = \frac{-b_\lambda\ -\ \sqrt{b_\lambda^2-4a_\lambda c_\lambda}}{2a_\lambda}.
$$

We then construct ω_ℓ using the definition in proposition 2.

A.11 Derivation of equation (26)

Starting from:

$$
\pi_t = Amc_t + Bmc_{t-1} + C\pi_{t-1} + D\pi_{t-2}
$$

where the coefficients A, B, C, D satisfy the restrictions from the proof of proposition 4 in section A.8. Rewrite the dynamics of the price level as:

$$
p_t\ =\ a_0\, mc_t^n\ +\ a_1\, mc_{t-1}^n\ +\ \phi_1 p_{t-1}\ +\ \phi_2 p_{t-2}\ +\ \phi_3 p_{t-3}
$$

where:

$$
\begin{aligned}
\Delta &\equiv \gamma\vartheta-\rho\beta+\kappa(1-\Omega)\Big(1-\frac{\lambda}{\rho}\Big),\\
a_0 &\equiv \frac{\kappa(1-\Omega)(1-\lambda/\rho)}{\Delta},\\
a_1 &\equiv -\gamma\, a_0,\\
\phi_1 &= \frac{(\gamma\vartheta-\rho\beta)\Big(1+\lambda+\frac{1}{\vartheta}\Big)+\gamma\,\kappa(1-\Omega)\Big(1-\frac{\lambda}{\rho}\Big)}{\Delta},\\
\phi_2 &\equiv -\frac{(\gamma\vartheta-\rho\beta)\Big(\lambda+\frac{\lambda}{\vartheta}+\frac{1}{\vartheta}\Big)}{\Delta},\\
\phi_3 &\equiv \frac{(\gamma\vartheta-\rho\beta)\Big(\frac{\lambda}{\vartheta}\Big)}{\Delta}.
\end{aligned}
$$

We obtain a formula for year-over-year inflation in terms of the path of realized nominal marginal cost by repeatedly substituting the lags of the price:

$$p_t - p_{t-12} = \sum_{\tau=0}^{11} \psi_\tau \Big(a_0\, mc^n_{t-\tau} + a_1\, mc^n_{t-\tau-1} \Big) + (\psi_{12} - 1)\, p_{t-12} + \Big(\phi_2 \psi_{11} + \phi_3 \psi_{10} \Big) p_{t-13} + \phi_3 \psi_{11}\, p_{t-14}.$$

The coefficients are obtained from the recursion:

$$\psi_0 = 1, \qquad \psi_1 = \phi_1, \qquad \psi_2 = \phi_1^2 + \phi_2, \qquad \psi_\tau = \phi_1 \psi_{\tau-1} + \phi_2 \psi_{\tau-2} + \phi_3 \psi_{\tau-3} \quad (\tau \geq 3).$$

B Data sources and data cleaning

In this section we provide additional details on the various administrative data sources and data cleaning procedures that underlies the operationalization of our empirical framework. Some of the material in this section draws on Gagliardone et al. (2025, 2024), amended to account for the higher frequency of the estimation and additional data sources.

PRODCOM. We use information from PRODCOM to compute the monthly change in product- and firm-level prices and to define the boundaries of markets (industries) in which firms compete. PRODCOM — derived from the French PRODuction COMmunautaire — is a large-scale survey commissioned by Eurostat and administered in Belgium by the National Statistical Office. The PRODCOM sampling strategy is designed to cover at least 90% of domestic production value within each manufacturing industry (4-digit NACE codes) by surveying all firms operating in the country with (i) a minimum of 20 employees or (ii) total revenue above 4.5 million euros (European Commission 2014). Firms are required to disclose, on a monthly basis, product-specific physical quantities (e.g., volume, kg., m^2, etc.) of production sold and the value of production sold (in euros) for all their manufacturing products. Products are defined in PRODCOM by an 8-digit PC code (e.g., 15.20.13.61 is "Mens's sandals", 15.20.13.62 is "Womans's sandals", and 15.20.13.63 is "Children's sandals"). Industries are defined by the first four digits of the product code (e.g., 15.20 is "Manufacture of footwear"). Sectors, i.e. more aggregate structures within the economy, are defined by the first two digits of the product codes (e.g., CB is "Textile, apparel, leather and related products").

In the raw data, there are approximately 4,000 product headings distributed across 13 manufacturing sectors. The PC product codes have been revised several times between 1995 and 2023, with a substantial overhaul in 2008. We use the conversion tables pro-

vided by Eurostat and firm-specific information on firms' product portfolios to harmonize the 8-digit product codes across consecutive months and harmonize 4-digit industry codes over time.[19] In most cases, the conversion tables provide a unique mapping of the 8-digit product codes across consecutive years. In a limited number of cases, the mapping is many-to-one, one-to-many, or many-to-many. The many-to-one mapping is straightforward, while the one-to-many and many-to-many mappings could be problematic. We are able to handle most of these cases using information on the basket of products produced by each firm.[20] In a limited number of cases, we do not have sufficient information to resolve the uncertainty regarding the mapping. We drop these observations from the sample.

We construct product-level prices (unit values) by dividing product-level sales by product-level quantities sold. As explained in the body of the paper, we are interested in domestic prices, i.e., prices charged by producers in Belgium. PRODCOM does not require firms to separately report production and sales to domestic and international customers. Therefore, we recover domestic values and quantities sold by combining information from PRODCOM with data on firms' product-level exports (quantities and sales) available through Belgian Customs (for extra-EU trade, "Extrastat") and the Intrastat Inquiry (for intra-EU trade).[21] We use the official conversion tables provided by Eurostat to map the CN product code classification used in the international trade data to the PRODCOM product code classification. In most cases, the CN-to-PC conversion involves either a one-to-one or many-to-one mapping, which poses no issues. We drop observations that involve one-to-many and many-to-many mappings.

We apply the following filters and data manipulations to the PRODCOM data set. First, we retain firms' observations if there was positive production reported for at least one product in all months throughout the year. In the rare cases where a firm reports positive values but quantities are missing, we impute the quantity sold from the average value-to-quantity ratio in the months where both values and quantities are reported. Second, we require firms to file VAT declarations and Social Security declarations (as explained below). These two data sources allow us to measure firms' marginal costs.

[19] The official conversion tables are available at the Eurostat webpage. The harmonization of the industry code essentially consists of harmonizing the NACE Rev.1 industry, used before 2008, to the NACE Rev.2 industry codes, used from 2008.

[20] For example, consider a case where the official mapping indicates that product 11.11.11.11 in year t could map to either 22.22.22.21 or 22.22.22.22 in year $t+1$. Suppose two firms, f_1 and f_2, report in period t sales of product 11.11.11.11 in year t. If f_1 reports *only* sales of 22.22.22.21 and f_2 *only* reports sales of 22.22.22.22 in year $t+1$ we infer that we should map 11.11.11.11 to 22.22.22.21 for the former and 11.11.11.11 to 22.22.22.22 for the latter.

[21] In constructing our measure of domestic sales, we remove transit trade, i.e. imported goods that briefly pass through the Belgian territory and are re-exported without Belgian ownership.

Value added tax and social security declarations. We leverage data from two administrative sources to measure firms' total production (turnover) and variable production costs on a monthly basis. First, Belgian firms file VAT declarations to the tax authority that contain information on the total sales of the enterprise as well as information on purchases of raw materials and other goods and services that entail VAT-liable transactions, including domestic and international transactions. While this submission requirement applies to quasi all Belgian firms, small firms — i.e. firms having an annual turnover < 2.5M euros excl. VAT — are allowed to submit the information on a quarterly frequency. We restrict our analysis to the (larger) firms that submit on a monthly frequency. In view of the the stratified sampling strategy of PRODCOM (supra), this has a small impact on the composition of our sample.

Second, we obtain information on employment and labor costs (wage bill) from the Social Security declarations filed quarterly by each Belgian firm with the Department of Social Security in Belgium. We smooth the quarterly wage mass equally across the three months of the quarter (and assume an equal headcount in each month). We qualify this as an innocuous assumption in view of tight labor market protection laws in Belgium, such as restrictions on hiring or firing, and work rules. Moreover, in manufacturing, the segment of a firms' labor force that can most easily be scaled up/down in the short run is often employed through temporary employment agencies. Importantly, in such a contracting setup, the wage expense shows up as a component of the monthly VAT declaration of the manufacturing firm (as a payment to the employment agency, who, in turn pays a wage that is tracked through Social Security declarations).

We sum firm-monthly level expenses on intermediates and labor to obtain a measure of total variable costs, which we use to construct firms' marginal costs. We multiply these costs by the ratio of total manufacturing sales (from PRODCOM) to total sales (from the VAT declarations) to adjust for the fact that some firms also have production outside manufacturing.[22]

Data filters. To arrive at our baseline sample, we apply the following data-cleaning steps to address missing values and outliers. (i) We focus on manufacturing industries defined by the NACE 2-digit codes 10—30, dropping from our sample all product headings that correspond to mining and quarrying, and all product codes corresponding to industrial services. (ii) As is standard, we exclude firms that operate in the "Coke and refined petroleum products" sector and the "Pharmaceuticals, medicinal chemical, and botanical products" sector, whose output prices prices are often determined in international (spot) markets or

[22] We drop observations referring to firms whose sales from manufacturing products (as measured in PRODCOM) are lower than seventy percent of total firm-level sales (as reported in the VAT declarations). This ensures that our sample includes firms whose real activity is primarily, if not entirely, in manufacturing.

privately bargained. We also exclude firms operating in the "Other manufacturing and repair and installation of machinery and equipment" sector, a residual grouping that encompasses firms producing diverse and varied products for which it is difficult to define an appropriate set of competitors. Table A.1 reports the final list of manufacturing sectors and their 2-digit PC codes. (iii) We keep only observations for which we are able to compute product-level price indexes, the corresponding quantity indexes, and marginal costs. (iv) We drop observations for which the month-to-month change of either the firm-level price index or marginal costs is greater than 100% in absolute value. (v) Finally, for each firm-industry pair that enters our dataset discontinuously, we keep only the longest continuous time spell. This ensures that each time series used in the estimation has no gaps. We only preserve firm-industry observations for which the time spell exceeds 24 months.

Business survey of the National Bank of Belgium. We merge the baseline sample with the NBB-BS survey after performing two data adjustments. First, since the NBB-BS survey records responses at the firm–product level, we aggregate these to the firm–industry level by taking the mode across products within each industry (affecting approximately 0.81% of observations). Second, although the survey tracks firms consistently over time, occasional non-reporting or data quality issues create gaps in the panel. We address these rare cases by linearly interpolating the missing values, allowing us to construct the longest possible time series for each firm.

Table A.1: List of manufacturing sectors

Sector	Sector definition	NACE 2-digits codes
CA	Food products, beverages and tobacco products	10–12
CB	Textiles, apparel, leather and related products	13–15
CC	Wood and paper products, and printing	16–18
CE	Chemicals and chemical products	20
CG	Rubber and plastics products, and other non-metallic mineral products	22–23
CH	Basic metals and fabricated metal products, except machinery and equipment	24–25
CI	Computer, electronic and optical products	26
CJ	Electrical equipment	27
CK	Machinery and equipment n.e.c.	28
CL	Transport equipment	29–30

Notes. This table reports the list of manufacturing sectors in our sample and the corresponding 2-digit NACE codes.

B.1 Estimates of the returns to scale

Estimates are taken from Supplemental Appendix OA.1 in Gagliardone et al. (2025).

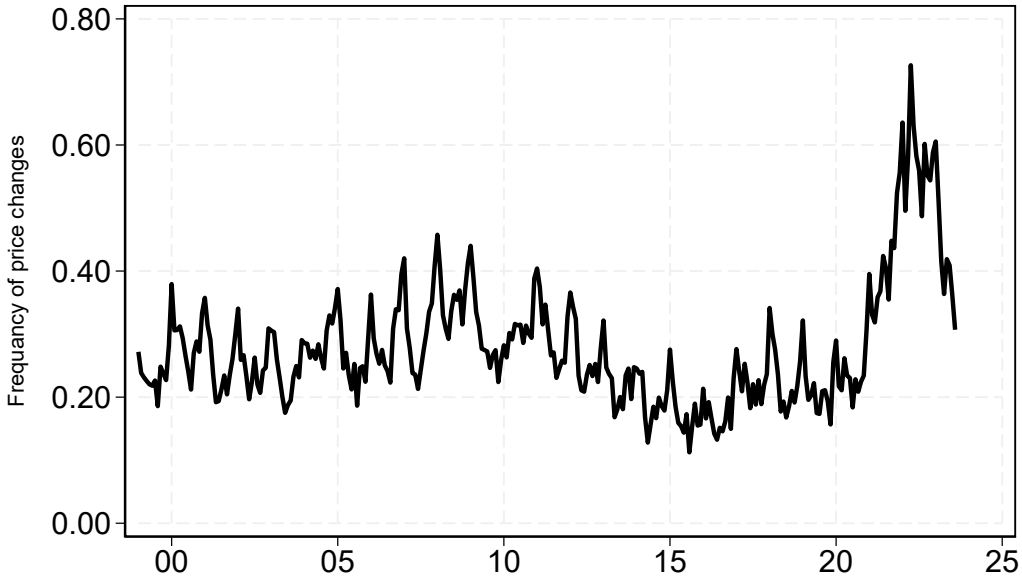
Table A.2: Estimates of output elasticities and returns to scale

Sector	Output elasticities			Returns to scale	
	Labor (ϑ^l)	Capital (ϑ^k)	Intermediates (ϑ^m)	Short-run ($\vartheta^l + \vartheta^m$)	Long-run ($\vartheta^l + \vartheta^m + \vartheta^k$)
CA	0.214	0.026	0.752	0.966	0.992
CB	0.200	0.028	0.764	0.964	0.992
CC	0.230	0.033	0.731	0.960	0.994
CE	0.217	0.039	0.720	0.937	0.976
CG	0.219	0.008	0.768	0.987	0.996
CH	0.188	0.028	0.776	0.963	0.991
CJ	0.177	0.036	0.782	0.960	0.996
CK	0.247	0.054	0.684	0.932	0.986
CL	0.243	0.036	0.725	0.968	1.004
Aggregate	0.212	0.026	0.753	0.965	0.992

Notes. This table reports the within-sector average production function elasticities of Gagliardone et al. (2025), estimated following the approach in Lenzu et al. (2025). The first column indicates the manufacturing sector. The subsequent three columns report the estimates obtained from a quantity production function estimation. The following two columns report the long-run and short-run returns to scale. Each row corresponds to a different manufacturing sector. The last row is a sales-weighted average of the sectoral estimates.

B.2 Plot of the frequency of price changes

Figure A.1: Monthly frequency of price changes



Notes. Time series of the frequency of price changes.