Markup Distribution and Aggregate Dynamics*

Mario Giarda — Central Bank of Chile
Will Jianyu Lu — Central Bank of Chile
Antonio Martner — UCLA & Central Bank of Chile

September, 2025

Abstract

Using administrative firm-level data from Chile, we estimate monthly firm-level markups and document that average markups are countercyclical, driven by firms at the top of the markup distribution. To interpret these facts, we develop a New Keynesian model with heterogeneous firms, endogenous markups, and nominal rigidities. Firms with greater market power adjust prices less, amplifying and prolonging output responses through persistent markup dispersion. This amplification is stronger under high-demand superelasticity and persistent shocks. Quantitatively, the aggregate GDP response to an aggregate demand shock can be up to 20% larger than in representative-firm models.

*We thank seminar participants at CREI, SED, Summer SED, CEBRA, LACEA, CEMLA, and Central Bank of Chile for excellent suggestions. The views expressed are those of the authors and do not necessarily represent the views of the Central Bank of Chile or its board members. This study was developed within the scope of the research agenda conducted by the Central Bank of Chile (CBC) in economic and financial affairs of its competence. The CBC has access to anonymized information from various public and private entities, by virtue of collaboration agreements signed with these institutions. To secure the privacy of workers and firms, the CBC mandates that the development, extraction and publication of the results should not allow the identification, directly or indirectly, of natural or legal persons. Officials of the Central Bank of Chile processed the disaggregated data. All the analysis was implemented by the authors and did not involve nor compromise the Servicios de Impuestos Internos de Chile. The information contained in the databases of the Chilean IRS is of a tax nature originating in self-declarations of taxpayers presented to the Service; therefore, the veracity of the data is not the responsibility of the Service. Authors emails: mgiarda@bcentral.cl, wlu@bcentral.cl, amartner@ucla.edu.

1 Introduction

Market power and nominal rigidities generate firm-level wedges that distort responses to macroeconomic shocks. In New Keynesian models, these wedges—markups—arise from monopolistic competition and price stickiness, and give rise to the New Keynesian Phillips Curve (Galí, 2015). This curve typically relates price inflation to measures of economic slack, such as the output gap or marginal costs (Gagliardone et al., 2023). However, most applications abstract from the micro-level origins of aggregate dynamics, overlooking heterogeneity in market power and price adjustment costs. As a result, they miss how systematic differences in markups and price-setting behavior across firms can shape not only inflation but also the responses of aggregate output and consumption to shocks.

Recent research shows that firm-level distortions can aggregate in nonlinear ways, implying that the distribution of markups, not just specific moments, matters for macroeconomic dynamics (Baqaee et al., 2024; Kouvavas et al., 2021). Yet the implications of markup heterogeneity for inflation and real activity remain underexplored in models with nominal rigidities. In particular, when demand shifts toward firms with higher markups and sluggish price adjustment, inflationary pressures may be muted, while real output and consumption can react more sharply. Ignoring this mechanism risks mischaracterizing the transmission channels of monetary and other aggregate demand shocks.

In this paper, we study how firm-level markup heterogeneity shapes the transmission of aggregate demand shocks to aggregate inflation and output in the presence of price rigidities. Following the aggregate version presented in Galí (2015), we write a firm-level Phillips curve expressed in terms of markup gaps (the difference between actual markup and a desired–flexible-price– markup), and show that the level of markup interacts with the markup response to a shock. We also uncover that after aggregation, aggregate inflation depends on both the average markup and the cross-sectional covariance between markup responses and markup levels, and show that shocks are amplified or dampened depending on this covariance.

Then, we test the predictions of the model on the data. We use monthly administrative tax data from Chile and document that markups are countercyclical on average, and that firms with higher markups exhibit the strongest countercyclical behavior. We show that this countercyclicality holds both unconditionally and conditional on demand shocks. This implies that the covariance between levels and responses is negative, generating stronger nominal rigidities. Motivated by these patterns, we develop a general equilibrium model with endogenous markups and price rigidities arising from Rotemberg (1982) price adjustment costs and Kimball (1995) demand, following Edmond et al. (2023). We calibrate the three main parameters with the data and find an amplification of demand shocks in our model due to a persistent rise in markup dispersion.

We begin by presenting a New Keynesian Phillips Curve (NKPC) where aggregate inflation depends on the distribution of firm-level markup gaps, the difference between observed markups under sticky prices and desired markups under flexible prices. We first show analytically that the firm-level Phillips curve can be expressed in terms of markups, which can be measured directly in the data, rather than relying on a more indirect measure, the output gap. We find that when aggregating these Phillips curves, a new channel for inflation dynamics emerges: the covariance between fluctuations in markup gaps and the markup level.

Assuming Rotemberg price adjustment costs and isoelastic demand, we show that firms with higher markups endogenously face flatter Phillips curves and adjust prices less, despite identical adjustment costs. In response to an expansionary demand shock, demand shifts toward these initially high-markup firms. Because their markups respond more than their prices, this reallocation dampens the effect of the shock on aggregate inflation. Thus, pricing heterogeneity has first-order implications for the transmission of demand shocks and overall macroeconomic dynamics.

Then, we use administrative tax data from Chile, which cover the universe of the formal economy, to estimate monthly firm-level markups based on quantity rather than expenditure production functions, following De Loecker and Warzynski (2012). We find that the mean markup is countercyclical, increasing during economic downturns and declining during expansions. This countercyclicality is not uniform across the markup distribution; firms at the top of the distribution exhibit markups with stronger countercyclical behavior compared to firms at the bottom. In addition, we document a negative correlation between markup growth and both GDP growth and inflation.

We employ panel local projections to study the response of firm-level markups to aggregate demand shocks, using an identified monetary policy shock from Aruoba et al. (2021) and an identified foreign financial shock, the Excess Bond Premium by Gilchrist and Zakrajšek (2012).[1] We find that markups are, on average, countercyclical in response to both shocks, and they remain above their baseline levels for at least two years after the

[1] Aruoba et al. (2021) shows that in response to a contractionary monetary there is a decline in output and inflation, using their high frequency identification. Aldunate et al. (2025) show that an EBP shock by Gilchrist and Zakrajšek (2012) also generates the same output and inflation comovement, and thus, looks like a demand shock.

shock. These patterns indicate that firm-level price adjustment is limited in the short run, making markup variation a key margin of adjustment over the business cycle, consistent with New Keynesian models' predictions under price rigidities.

Using the local projection specification, we study the heterogeneous response of markups to aggregate demand shocks. Based on Ottonello and Winberry (2020), we compare firms' responses above the mean markup to those at the mean. We find that high-markup firms exhibit significantly stronger markup responses to both monetary policy and foreign financial shocks. These effects are persistent, lasting for more than two years. This persistent widening of markup differentials indicates that aggregate demand shocks affect not only the level of aggregate markups but also their distribution across firms. Our results reveal an important asymmetry: firms with greater market power absorb shocks primarily through markup adjustments, without changing their prices, and absorb marginal cost changes, whereas firms below the aggregate mean markup adjust their prices, maintaining their markup levels relatively unchanged.

To study the extent to which these results have effects at the aggregate level, we build a general equilibrium model in which firms differ in productivity and market power, and face quadratic price adjustment costs, following Rotemberg (1982). The model embeds a Kimball (1995) demand aggregator, implying that demand elasticities vary endogenously with firms' market shares. This structure generates endogenous and time-varying desired markups that increase with firm size. Combined with nominal rigidities, this heterogeneity gives rise to firm-specific Phillips curves with slopes that decline in firm markups.

High-markup firms, facing relatively more inelastic demand and flatter Phillips curves, adjust prices less in response to shocks and absorb fluctuations through markup variation. As a result, price rigidity is not an exogenous friction, but an outcome of market structure and firm heterogeneity. We find that with the baseline calibration, the cumulative effect of a demand shock is about 20% larger than in its CES counterpart. However, this result depends on the persistence of the shock. When a shock is more persistent, the effects are amplified even more. The reason is that in response to a shock that generates higher dispersion in markups, subsequent shocks affect a more dispersed distribution, generating a stronger response. Hence, mechanisms that generate more persistent effects of the shock amplify the effects of markup dispersion.

Related Literature. Our paper relates to three strands of literature on markups and macroeconomic dynamics. The first examines how aggregate inefficiencies shape inflation, output, and consumption, including Galí et al. (2007), Bils et al. (2018), and Nekarda and Ramey (2020). These studies focus on average distortions in New Keynesian models but abstract from firm-level heterogeneity in markup behavior.

The second strand explores unconditional markup cyclicality, including Hong (2017), Afrouzi and Caloi (2023), Anderson et al. (2018), and Burstein et al. (2020). These papers document how markups comove with the business cycle without isolating specific aggregate shocks. Burstein et al. (2020), in particular, show how markup heterogeneity in a granular oligopolistic model generates aggregate volatility, and suggest enriching such models with price-setting frictions and demand shocks—directions our paper explicitly pursues.

The third strand analyzes heterogeneous firm-level markup responses to identified policy shocks. Meier and Reinelt (2022) emphasize the role of nominal rigidity in driving markup dispersion and misallocation effects. Chiavari et al. (2021) study heterogeneity by firm age using a model close to ours. Kouvavas et al. (2021) and Höynck et al. (2023) study how the transmission of monetary policy varies with market power but focus on sectoral or aggregate data. In contrast, we document how markup levels and their dynamic responses interact with firm-specific pricing frictions, shaping aggregate inflation and output.

2 Markup Heterogeneity in the New Keynesian Model

This section develops a framework that connects inflation dynamics to the distribution of firm-level markups in the presence of nominal rigidities. We show that aggregate inflation can be expressed as a function of the markup gap—the difference between a firm's observed sticky-price markup and its desired flexible-price markup—and the distribution of firm-level markups. We find that the average markup gap and the covariance of markup gaps with slope parameter (which captures the sensitivity of inflation to markups at a firm level) determines the response of aggregate inflation to macroeconomic shocks. This formulation offers an alternative to output-gap–based inflation models and yields testable implications regarding the role of firm-level markup heterogeneity.

Markups in the Aggregate New Keynesian Phillips Curve (NKPC). Following Galí (2015), in a setting with nominal rigidities, monopolistic competition in intermediate goods, and market clearing in factor inputs, the NKPC is the relationship between output gap and aggregate inflation:

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

where π_t denotes aggregate inflation, $\tilde{y}_t = y_t - \overline{y}_t$ is the output gap, ζ is the slope of the Phillips curve, and β is the discount factor. Natural output $\overline{y}_t$ is defined as the output level that would prevail without nominal rigidities.

Empirically, estimating $\tilde{y}_t$ poses practical challenges. A common approach to estimate the output gap decomposes output into cyclical and trend components using filters, assuming that the trend approximates natural output. However, this method is sensitive to assumptions about the underlying process and may include structural inflation drivers like persistent demand shocks. Moreover, the derivation of Equation (1) assumes a strong link between the output gap and marginal costs. This relationship breaks down with frictions in labor or input markets, which is why empirical estimates of ζ are typically very low. (see Gagliardone et al., 2023)

An alternative formulation expresses inflation directly as a function of real marginal costs, yielding a more primitive version of the NKPC:

$$\pi_t = \lambda mc_t^r + \beta \mathbb{E}_t\{\pi_{t+1}\},$$

where $mc_t^r = mc_t - p_t$ is real marginal cost, and λ is a slope parameter. For firms facing isoelastic demand, the real marginal cost is proportional to the inverse of the markup. This implies that marginal cost–based inflation equations, such as the NKPC, can equivalently be expressed in terms of markups.

We define the markup gap as the deviation of the observed (sticky-price) markup from the desired (flexible-price) markup: $\tilde{\mu}_t = \log \mu_t - \log \mu_t^d$. This substitution yields an equivalent NKPC in terms of markup gaps:

$$\pi_t = -\lambda \tilde{\mu}_t + \beta \mathbb{E}_t\{\pi_{t+1}\}. \tag{2}$$

This expression shows that deviations of observed markups from their desired levels generate inflationary pressures in the presence of price rigidities. This a generalization of expressions derived in Galí (2015) to endogenous and time-varying desired markups.[2] This

[2]The basic New Keynesian model in Galí (2015) assumes monopolistic competition, which has a constant and homogeneous desired markup that disappears from the log-linearized New Keynesian Phillips curve. We relax that assumption and let μ_t^d be time-varying.

expression provides a measurable and theoretically consistent alternative to the output gap as a measure of economic slack.

Heterogeneity in the NKPC. We extend Equation (2) to a heterogeneous firm environment where markups and slope parameters vary across firms. At the firm level, the inflation equation becomes:

$$dp_{it} = -\lambda_i \tilde{\mu}_{it} + \beta \mathbb{E}_t\{dp_{it+1}\}, \tag{3}$$

where dp_{it} denotes the log change in the price of firm i, $\tilde{\mu}_{it} = \log \mu_{it} - \log \mu_{it}^d$ is the firm-specific markup gap in logarithm, and λ_i captures firm-specific slope coefficients that depend on pricing frictions and demand elasticities.

Let α_{it} be firm i's final demand Domar weight defined as each firm's expenditure share in the target price index (e.g., CPI or GDP deflator), so that aggregate inflation is the α_{it}-weighted mean of firm-level price changes. . Aggregating Equation (3) over a continuum of firms indexed by $i \in [0, 1]$, we obtain:

$$\pi_t = -\int_0^1 \alpha_{it} \lambda_i \tilde{\mu}_{it} \, di + \beta \mathbb{E}_t\{\pi_{t+1}\}. \tag{4}$$

If there is not firm heterogeneity in the slope ($\lambda_i = \lambda$ for all i), this reduces to equation (2), with $\tilde{\mu}_t = \int_0^1 \alpha_{it} \tilde{\mu}_{it} \, di$ and $\pi_t = \int_0^1 \alpha_{it} dp_{it} \, di$. More generally, heterogeneity in λ_i introduces an additional term capturing the interaction between markup gaps and slope variation. We can rewrite Equation (4) as[3]:

$$\pi_t = -\left(\mathbb{E}_\alpha[\lambda_i] \cdot \mathbb{E}_\alpha[\tilde{\mu}_{it}] + \mathrm{Cov}_\alpha(\lambda_i, \tilde{\mu}_{it})\right) + \beta \mathbb{E}_t\{\pi_{t+1}\}. \tag{5}$$

Equation (5) decomposes inflation dynamics into three channels. First, the usual expected inflation channel in NKPC, and two channels depending on markups. A first term captures the weighted average effect of firm-level markup gaps on inflation, scaled by the average slope. The second term reflects the covariance between firm-level NKPC slopes and markup gaps. When this covariance is negative, it countervails the first term, and inflation is dampened. To explain this result further, we provide an example in which we can express the slope λ_i explicitly depending on markups, and thus, we can study the

[3] Applying the identity $\int X_i Y_i di = \mathbb{E}_\alpha[X_i]\mathbb{E}_\alpha[Y_i] + \mathrm{Cov}_\alpha(X_i, Y_i)$, with the subscript α denoting the moment using the measure α.

sign of $\text{Cov}_\alpha(\lambda_i, \tilde{\mu}_{it})$.

Example: Rotemberg Price Adjustment Costs. Consider the case of quadratic price adjustment costs following Rotemberg (1982), where each firm maximizes lifetime profits with a common price adjustment parameter θ. Suppose that firm i faces isoelastic demand with elasticity σ_i. As we show in Appendix C, the log-linearized firm-level NKPC has a slope given by

$$\lambda_i = \frac{\sigma_i - 1}{\theta}.$$

For profit-maximizing firms under isoelastic demand, the elasticity σ_i can be expressed in terms of the markup: $\sigma_i = \frac{\mu_i}{\mu_i - 1}$. Substituting, we obtain:

$$\lambda_i = \frac{1}{(\mu_i - 1)\theta}.$$

Replacing λ_i in equation (3) gives the firm-level pricing equation:

$$dp_{it} = -\frac{1}{(\mu_i - 1)\theta}\tilde{\mu}_{it} + \beta\mathbb{E}_t\{dp_{it+1}\}. \tag{6}$$

Equation (6) highlights that even with a common price adjustment cost θ, nominal rigidities vary across firms due to differences in their markup levels. Firms with higher markups face a flatter Phillips curve and are, as a consequence, subject to stronger nominal rigidities. This implies that, in response to shocks, these firms adjust their markups more when marginal costs change.

The intuition of the previous result is as follows. When facing a demand shock that raises marginal costs, high-markup firms—which face low demand elasticity—have less incentive to adjust their prices than low-markup firms. This occurs because high-markup firms lose only a small portion of their profits from not adjusting, and this loss may be smaller than the cost of price adjustment. As a result, they may choose not to change their prices and adjust their markups instead. Consequently, firms with higher markups exhibit stronger markup responses compared to firms with lower markups. As shown in Equation (6), both the level and the change in markups interact within the Phillips curve, resulting in heterogeneous slopes and a negative covariance between markup levels and their fluctuations. We study the empirical relevance of this covariance in the next section.

To show the effects of heterogeneous markups on aggregate inflation, we aggregate

Equation (6) across firms and applying the definition of covariance, we obtain:

$$\pi_t = -\frac{1}{\theta}\left(\mathbb{E}_\alpha\left[\frac{1}{\mu_i - 1}\right] \cdot \mathbb{E}_\alpha[\tilde{\mu}_{it}] + \text{Cov}_\alpha\left(\frac{1}{\mu_i - 1}, \tilde{\mu}_{it}\right)\right) + \beta\mathbb{E}_t\{\pi_{t+1}\}. \tag{7}$$

Equation (7) shows that with Rotemberg, inflation dynamics depends on the average markup gap and the covariance between firms' steady-state markup levels and their markup responses. When this covariance is negative, aggregate inflation is smaller than a non-price rigidity economy. Nominal rigidities emerge endogenously from the distribution of firm-level markups, even when adjustment costs are homogeneous.

This framework is robust to both supply and demand shocks. However, the effects of supply shocks on marginal costs are ambiguous. For example, a negative aggregate productivity shock raises input prices while reducing output, leading to offsetting movements in marginal cost components. Consequently, supply shocks are ill-suited for analyzing the cyclicality of markups under price rigidity. In contrast, demand shocks—such as a contractionary monetary policy—move prices and quantities in the same direction, generating a clearer link between marginal costs and markups when prices are sticky.

These unambiguous predictions make demand shocks a more suitable setting to test the implications of expressions like Equation (6), which predict a negative relationship between markups, aggregate output, and prices. Moreover, demand-driven fluctuations allow us to assess whether this relationship is stronger for firms with higher average markups, thereby providing empirical support for the role of heterogeneity in the dynamics described by Equation (6).

3 Markup Estimation and Cyclical Behavior

This section examines the relationship between firm-level markups and the business cycle using detailed administrative tax data from Chile, covering over 6 million firm-month observations across 626 industries. Relying on the responses of firm-level markups to identified demand shocks—estimated through local projections using monetary policy surprises and shocks to financial conditions—we find countercyclical markups. This result is particularly pronounced among high-markup firms; i.e., firms at the top of the markup distribution exhibit the strongest countercyclical behavior.

3.1 Data and markup estimation

Data. We use firm-level monthly data from different Chilean Internal Revenue Service (SII) sources. The Value Added Tax monthly form (Form 29) provides information on firm sales, materials expenditure, and investment. Monthly labor (headcounts) and wage data come from a separate SII form (DJ1887). Data on capital stock is sourced from the annual income tax form (Form 22). Using the perpetual inventory method, we combine the yearly capital stock with monthly investment data to construct firm-level monthly capital stock. The last source is the universe of electronic tax invoices available from 2014 onward, which provides product-level information, including prices and quantities, allowing us to construct variables for output and materials free from price variation.

We construct firm-level price indices for both output and intermediate inputs to separate quantity from price effects. We observe all goods sold to other firms and all intermediate inputs purchased for every formal firm in Chile. We generate Tornqvist quantity indices for production and intermediate inputs using this comprehensive invoice-level data.[4] We selected 2014 as our base year since it marks the beginning of our price data for firm-to-firm transactions. This approach aligns with established methods for estimating production functions at the firm level when price data is available (Dhyne et al. (2022) and De Roux et al. (2021)).

Specifically, we compute firm-specific annual weighted average prices (P_{igt}) for each product (g) sold by firm i during year t. We then construct firm-specific price indices (ΔP_{it}) for products observed in consecutive years:

$$\Delta \log P_{it} = \sum_{g} \frac{s_{igt} + s_{igt-1}}{2} \, \Delta \log(P_{igt}),$$

where s_{igt} represents the revenue share of product g for firm i at time t. We apply an analogous procedure to materials, ensuring output and input measures are free from price variation.

To ensure data quality, we perform minimal cleaning and define a firm as a tax ID that simultaneously satisfies three criteria: (i) positive revenue and material expenditure every year, (ii) capital level of at least USD 10, and (iii) at least one employee. Under this cleaning, we keep 23% of tax IDs, which we define as firms, while keeping 85% of the value added to the economy. Table 4 in Appendix A shows the median of main produc-

[4] We assume that the price index of a given seller-product is the same for final consumers and other firms.

tion variables for 2018 separately for 11 industries.

Markup estimation. To describe markup effects on inflation dynamics, we measure firm-level markups following the production function approach of De Loecker and Warzynski (2012). We estimate quantity-based rather than revenue-based production functions at the 6-digit industry level (626 industries in our dataset) using the Ackerberg et al. (2015) estimation method. We assume a Cobb-Douglas functional form for the quantity production function with three factors: labor, capital, and materials. The method uses materials' first-order condition to recover the variable input-output elasticity.

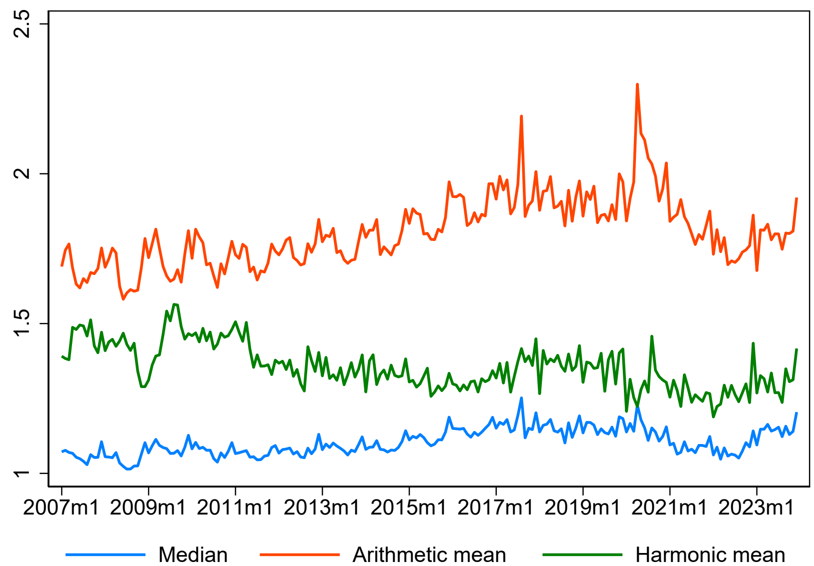
The production-based approach to measuring markups has a rich history, beginning with the seminal work of Hall (1988) and later expanded by De Loecker and Warzynski (2012). A central challenge in this literature concerns the distinction between revenue ($P \cdot Q$) and physical output (Q) when estimating production functions. Bond et al. (2021) raises concerns that revenue-based approaches may produce biased markup estimates. In contrast, De Loecker (2021) counters that the De Loecker and Warzynski (2012) methodology effectively handles price effects by treating them as omitted variables. De Ridder et al. (2022) offers a middle ground, suggesting that revenue-based methods yield reliable estimates of markup changes over time, even if the absolute markup levels contain bias. Our quantity-based approach circumvents many of these concerns by recovering a measure of physical output.

Our analysis of markup dynamics requires accurate markup levels and changes over time. While we do not intend to contribute to the markup estimation debate, we leverage our dataset to use the production approach framework. An advantage of our data is that we observe product-level prices, allowing us to estimate production functions based on quantity measures for both output and material usage rather than revenue and material expenditures.

For this section, we remain agnostic about the underlying market conduct that generates markups and focus solely on its measurement. To recover the markup, we calculate the ratio of material-output elasticity to the material revenue share.

We estimate sector-specific yearly production functions to recover material-output elasticities, which we assume to be time-invariant. Given our monthly data frequency, we compute monthly material revenue shares and construct monthly markups for every firm in our sample. Figure 1 illustrates that markups exhibit seasonality throughout 2007-2023 but have remained relatively stable using different firm-level weighting schemes.

Figure 1: Markup moments monthly evolution



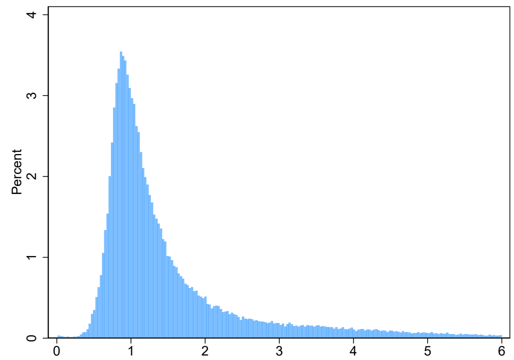
Notes: This figure displays the monthly evolution of firm-level markups in Chile from 2007 to 2023, computed using a production function approach based on physical quantities. Markups are calculated as the ratio between the material-output elasticity (estimated at the 6-digit industry level) and the firm's observed material revenue share. We show two aggregation methods: the harmonic mean, which gives more weight to firms with higher cost shares and accounts for firm-level heterogeneity in cost structure, and the arithmetic mean, which gives equal weight to all firms regardless of size or cost structure. Both series exhibit clear seasonal fluctuations but no strong long-run trend. The figure is based on over 6 million firm-month observations across all sectors of the Chilean economy.

A Measure of Permanent Markup. To investigate how the markup-GDP relationship varies across the distribution of markups, we first construct a measure of each firm's permanent markup level. We estimate the following fixed-effect regression to obtain measures of average markups by firm:

$$\log \mu_{it} = \alpha + \theta_i + \theta_{ms} + \theta_{ys} + \upsilon_{it}, \tag{8}$$

where α is the mean log of markup across all firms, θ_i is a firm fixed effect, θ_{ms} is a sector-month fixed effect, and θ_{ys} is a sector-specific trend. We define a firm's permanent markup as $\mu_i^P = \exp(\alpha + \theta_i)$ and plot its distribution in Figure 2. The distribution resembles a Pareto distribution with a high concentration of values between 0.5 and 2, with a rapidly decreasing probability for larger values.

Figure 2: Permanent markup μ_i^p distribution



Notes: This histogram shows the distribution of firm-specific permanent markups μ_i^p, constructed from firm fixed effects estimated in a panel regression of log markups on firm, sector-month, and sector-specific time fixed effects. The distribution is truncated at the 99th percentile to improve readability and ensure anonymity. Most firms exhibit markups between 0.5 and 2, with a long right tail consistent with a Pareto-like distribution.

Understanding how firm-level markups respond to macroeconomic conditions is central to analyzing their role in inflation dynamics. We now turn to an examination of the cyclical properties of markups. This analysis provides insights into how markup variations correlate with broader aggregate economic fluctuations.

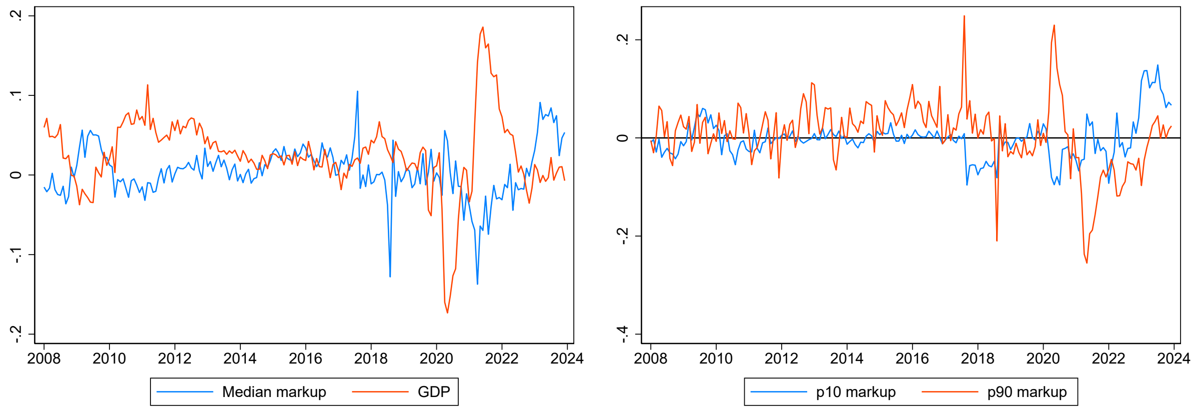
3.2 Markup Cyclicality

Now, we examine the markup cyclical properties. Growth in markups at the firm level is, on average, negatively correlated with both aggregate GDP and inflation. In addition, we find that this countercyclical pattern is primarily driven by firms at the top of the markup distribution, suggesting heterogeneity in how firms adjust pricing behavior to macroeconomic conditions.

Figure 3: Markup Cyclicality and Moments of its Distribution. 12 month log change

(a) Median Markup and GDP

(b) Markup P10 and P90



Notes: This figure explores the relationship between firm-level markup growth and aggregate economic activity in Chile, using 12-month log changes. Panel (a) shows the correlation between the median firm-level y-o-y markup growth and the y-o-y growth of Central Bank of Chile's monthly GDP index (IMACEC). Panel (b) disaggregates the analysis by plotting the 10th and 90th percentiles of the markup distribution. Markups are computed monthly from physical quantity-based production functions using firm-level tax and invoice data.

We begin by comparing 12-month log changes between measures of aggregate markups and GDP growth.[5] Figure 3a shows a negative correlation between the median markup and GDP with a correlation coefficient of -0.44, though this relationship occasionally decouples. The 2020-2022 period illustrates the countercyclical nature of markups: when GDP decreased during the COVID-19 pandemic and then strongly recovered by mid-2021, the median markup moved in the opposite direction.

This countercyclicality is not uniform across the markup distribution. As shown in Figure 3b, firms at the top of the distribution (90th percentile) exhibit stronger countercyclical behavior (correlation coefficient of -0.52) compared to firms at the bottom (10th percentile, correlation coefficient of 0.01).

The previous evidence is subject to aggregation bias and might hide the actual cyclicality of firm-level markups. To study the cyclicality further, we use panel regressions to evaluate the cyclicality of markup growth. In addition, we study the correlation of markup growth with aggregate inflation to test to what extent firms transfer aggregate inflation to prices and document whether firms absorb higher overall prices through markups. Negative correlations of markup growth with these variables are evidence of

[5]We use the Central Bank of Chile IMACEC, the GDP series monthly in Chile.

substantial price rigidities at a firm level.[6] To explore that, we run the following specification:

$$d \log \mu_{it} = \alpha + \beta X_t + \gamma \left(\mu_i^p - \overline{\mu}^p\right) X_t + \delta' Z_{it} + \theta_{im} + \epsilon_{it}, \tag{9}$$

where $d \log \mu_{it}$ is firm i's one-month change in the log markup, X_t can be either GDP growth or inflation, $(\mu_i^p - \overline{\mu}^p)$ is the distance of firm i permanent markup μ_i^p to the cross-sectional average of the permanent markup $\overline{\mu}^p$. This difference is our measure of heterogeneity where μ_i^p is obtained from Equation (8) and its distribution described in Figure 2. θ_{im} is a firm-month interacted fixed effect to deal with firm-level seasonality, and Z_{it} is a vector of controls including the real exchange rate, and lags of $d \log \mu_{it}$ and $d \log GDP_t$. We use all controls for the regressions using X_t = GDP growth, while for regressions using X_t = inflation, we use real exchange rate control only.

Therefore β represents either the cyclicality of markups or the correlation with aggregate inflation, unconditionally, and γ represents the differential effect of X_t on $d \log \mu_{it}$ for a firm whose permanent markup is one unit above the cross-sectional average ($\mu_i^p - \overline{\mu}^p = 1$), compared to a firm with average permanent markup. This implies that the marginal effect of X_t on markup growth is:

$$\frac{\partial \, d \log \mu_{it}}{\partial X_t} = \beta + \gamma(\mu_i^p - \overline{\mu}^p).$$

In the case that markups are countercyclical,i.e., $\beta < 0$, a $\gamma < 0$ means that firms with higher permanent markups have more countercyclical markups than firms at the mean, which implies that firms with higher permanent markups adjust their prices less in response to aggregate shocks since they absorb shocks through markups more strongly. In other words, high permanent markup firms have stronger price rigidities. The interpretation is equivalent when we study $X_t = \pi_t$: negative β, and γ imply that firms do not pass inflation completely to their prices and firms with higher markups have a lower pass-through of aggregate inflation to prices.

Table 1 presents the results of regression in Equation (9), suggesting that markups are countercyclical for the average firm. A one percent increase in GDP is associated with a 0.17% decrease in the average firm markup. A one p.p. increase in inflation is associated with a 0.48% decrease in the average firm markup. Column (2) shows that

[6] And also would reject the hypothesis of greed-inflation, which was popular during COVID pandemic times.

the interaction term $(\mu_i^p - \overline{\mu}^p)X_t$ has a negative and statistically significant coefficient for both GDP growth and inflation: $\gamma = -0.158$ for $X_t = d\log GDP_t$; this implies that firms with higher permanent markups respond more countercyclically to GDP growth—that is, their markups fall more when GDP rises—relative to firms with average or below-average markups. Similarly, in column (4), the interaction term is even more negative and significant: $\gamma = -0.579$ when $X_t = \pi_t$, indicating that firms with higher permanent markups experience a much more substantial decline in markups in response to inflation. Thus, firms with a permanent markup one unit above the mean have twice the cyclicality of markups than firms at the mean.

Table 1: Markup Cyclicality Regressions

	$X_t : d\log GDP_t$		$X_t : \pi_t$	
	(1)	(2)	(3)	(4)
X_t	-0.165***	-0.174***	-0.482***	-0.491***
$(\mu_i^d - \overline{\mu}^d)X_t$		-0.158***		-0.579***
R2	0.53	0.54	0.21	0.21
N	3,859,834	3,859,834	5,882,804	5,882,804

*** $p < 0.001$

Notes: This table reports panel regression results estimating how firm-level markup growth correlate with aggregate GDP growth or inflation. The dependent variable is the monthly log change in firm markup, $d\log\mu_{it}$. Columns (1) and (2) use monthly GDP growth ($d\log GDP_t$) as the main explanatory variable; columns (3) and (4) use monthly inflation (π_t). Each specification includes firm-month fixed effects, lagged markup growth, and relevant macro controls: exchange rate for all regressions and lagged GDP growth only in GDP specifications. The variable μ_i^p denotes a firm's permanent markup; $(\mu_i^p - \overline{\mu}^p)X_t$ captures how the response varies across the markup distribution.

These results reveal a countercyclical pattern in firm-level markups, with stronger effects among firms with higher permanent markups. This heterogeneity shows that firms' pricing power shapes their response to economic fluctuations. While these correlations with GDP and inflation are informative, they reflect reduced-form relationships that combine various underlying macroeconomic forces, including demand and supply shocks.

Hence, to better isolate the demand-driven components of markup dynamics, we now analyze firm-level responses to two types of identified demand shocks: monetary policy surprises and financial shocks measured by the Excess Bond Premium (EBP). Both shocks are widely used in the literature as plausibly exogenous aggregate demand shocks, which help us provide evidence of the role of markup adjustment over the business cycle.

3.3 Markup Responses to Demand Shocks

Next, we take advantage of our panel structure to analyze the response of firm-level markups–and its heterogeneous response— to identified demand shocks. To do so, we estimate the response of markups to monetary policy and financial shocks using panel local projections. Our goal is threefold: first, study the hypothesis that demand shocks generate a countercyclical response of markups on average; second, study if different firms' markups respond differently depending on their permanent markup; and third, we test if the previous findings are robust to different demand shocks. As we show below, markups are countercyclical on average; firms with higher markups have more responsive markups, and these results are robust to the two demand shocks we present next.

Shocks. First, we rely on monetary policy surprises (MPS) from Aruoba et al. (2021). They compute a monetary policy surprise from the difference between the Monetary Policy Rate (MPR) set by the Central Bank of Chile's Board meeting on a particular date and the expected rate provided by analysts as collected by Bloomberg the day of the meeting. From this high-frequency difference, Aruoba et al. (2021) get a series of monetary surprises, which we show in Appendix A, Figure 10. Aruoba et al. (2021) show that these surprises generate the expected theoretical results of monetary policy surprises: a contractionary shock generates a fall in GDP and in the consumer price index.

Second, we follow Aldunate et al. (2025) by using a global financial shock as an instrument for the Chilean borrowing rates as another aggregate demand shock. They use the Excess Bond Premium (EBP) by Gilchrist and Zakrajšek (2012) and build a time series of shocks using a monthly structural VAR for the US, assuming EBP shocks affect economic activity with a lag, while financial variables can react contemporaneously. Thus, they clean the original EBP series by other factors driving its fluctuations, like output and other observable financial variables. We borrow this series of shocks and use it as a proxy for demand shocks in Chile. Aldunate et al. (2025) show that an adverse EBP shock in the US deteriorates financial conditions in Chile and generates a decline in both GDP and the consumer price index, affecting the Chilean economy as a demand shock. We depict the time series of these shocks in Appendix A, Figure 12.

Empirical Specification. Inspired by Ottonello and Winberry (2020), we run panel Local Projections to study the effects of demand shocks on firm-level markups and their heterogeneity and run the series of $h \in [0, H]$ regressions given by:

$$\log \mu_{it+h} - \log \mu_{it-1} = \alpha_h + \beta_h shock_t + \gamma_h \left(\mu_i^p - \overline{\mu}^p\right) shock_t + \Gamma_h' Z_{it} + \theta_{im} + \epsilon_{it+h}, \tag{10}$$

where $\log \mu_{it+h} - \log \mu_{it-1}$ corresponds to the cumulative percent change (or long difference) in the logarithm of firm i's markup from month $t-1$ to $t+h$. Z_{it} is a vector of firm-level and aggregate controls that, in the baseline specification, only include employment of firm i in t to account for firm i's own business cycle at the moment of the shock.[7] θ_{im} is a firm-month fixed effect denoted to account for firm-level markup seasonality and firm-specific effects.

We are interested in the series of coefficients $\{\beta_h\}_{h=0}^H$ and $\{\gamma_h\}_{h=0}^H$. The coefficient β_h captures the average effect of a shock in t ($shock_t$) on markups h months after the shock. The variable $(\mu_i^p - \overline{\mu}^p)$ is the distance of firm i permanent markup μ_i^p to the cross-sectional average of the permanent markup $\overline{\mu}^p$. We use this difference analogously as Equation (9). Different from Ottonello and Winberry (2020), we study the differential effect of heterogeneity in the cross-section of the average markup, while they study the deviation with respect to firms' own average.

Therefore, the coefficient γ_h captures the additional effect of the shock on firms one unit above the mean of the permanent markup conditional on demand shocks. A positive γ_h indicates that firms with markups above the average tend to increase their markups more strongly than firms below the average–i.e., firms in the upper tail of the permanent markup distribution respond more intensively to the shock. For both shocks, we estimate Equation (10) for the 2010-2019 period, so we exclude the Great Recession and COVID pandemic to focus on periods absent of significant macroeconomic disruptions.

Response of Firm-Level Markups to Demand Shocks. As explained above, contractionary demand shocks generate a decline in marginal costs, which, due to price rigidities, imply an increase in markups.[8]

Figure 4 presents the impulse responses of firm-level markups to identified demand shocks: monetary policy surprises (MPS) and Excess Bond Premium (EBP) shocks, re-

[7] In Appendix B, we show the results for these estimations controlling for macroeconomic variables: GDP, inflation, and real exchange rate. All next results will be robust in including these variables.

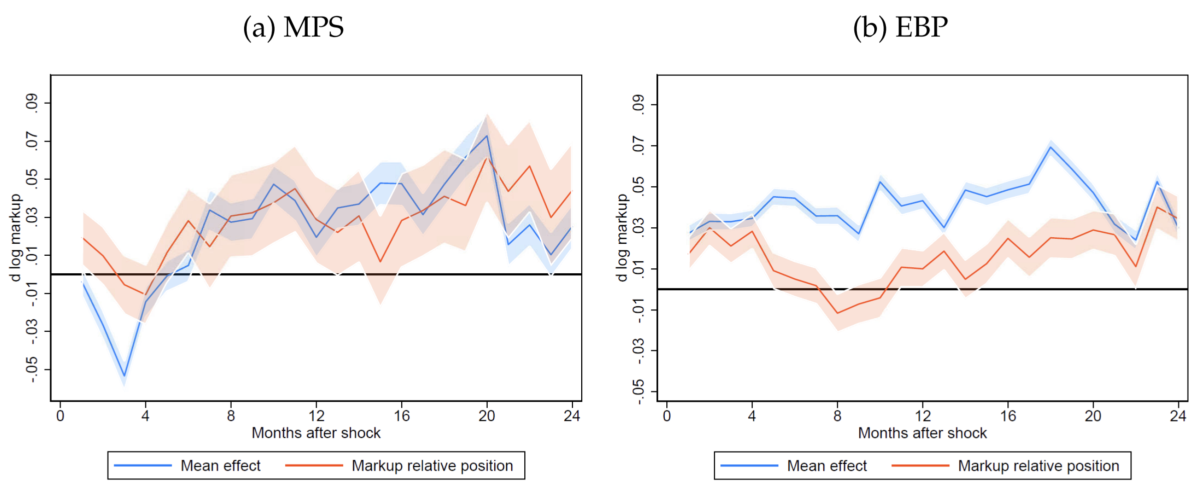
[8] This can be easily observed from the identity $P_t = \mu_t \times MC_t$, in which with P_t fixed, a decline in MC_t needs a rise in μ_t to keep satisfying the identity.

spectively. In both cases, the blue line shows the average cumulative markup response across all firms, while the red line captures the heterogeneous response by a firm's relative permanent markup position.

In response to a 1% monetary policy surprise, Figure 4a shows that average firm-level markups initially decline but then increase persistently, peaking around 6–7% by months 12–16 and remaining significantly elevated at the 24-month horizon. The interaction term is also strongly positive over time, indicating that firms with higher permanent markups exhibit a larger and more persistent increase in markups. This suggests that high-markup firms adjust prices less in response to monetary shocks, letting marginal costs fall while maintaining prices.

Figure 4b shows a qualitatively similar pattern for the case of an EBP shock. Following a shock to EBP that raises the corporate spread in Chile measured by the Corporate EMBI by 100 b.p., average markups increase by about 5–6% and remain persistently elevated. The relative markup response (red line) is again positive and statistically significant for much of the horizon, particularly beyond month 10. This result indicates that high-markup firms also disproportionately drive the increase in markups following financial shocks.

Figure 4: Response of Markups to Demand Shocks in the Baseline Specification.



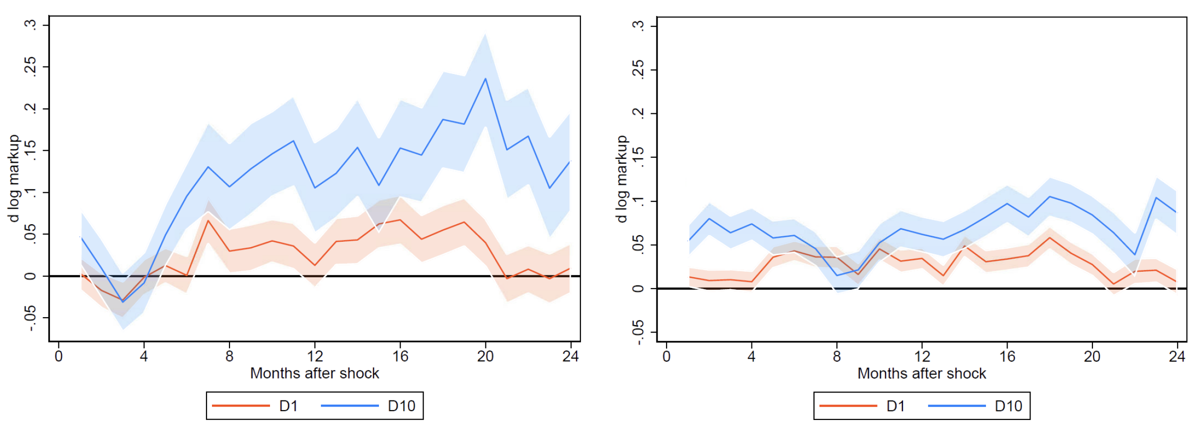
Notes: This figure presents the responses of firm-level markups to identified aggregate demand shocks using panel local projections over a 24-month horizon. Panel (a) shows the response to a 1% monetary policy surprise (MPS), and Panel (b) shows the response to an EBP shock which raises Chile's CEMBI by 100 b.p. The blue line plots the average markup response across all firms, while the red/orange line shows the additional response for firms whose permanent markup is one unit above the average.

Effects of demand shocks on markup dispersion. Our final empirical exercise analyzes the effect of demand shocks on the dispersion of firm-level markups. While previous figures show that average markups increase following contractionary monetary and financial shocks, we investigate whether the upper or lower tail of the markup distribution drives this increase. To do so, we estimate a variant of Equation (10) in which we interact the demand shock with indicator variables for the first (D1) and tenth (D10) deciles of the permanent markup distribution:

$$\log \mu_{it+h} - \log \mu_{it-1} = \alpha_0 + \beta_h^{D1} \cdot shock_t + \beta_h^{D10} \cdot shock_t + \Gamma_h' X_{it} + \alpha_{im} + \epsilon_{it+h}. \tag{11}$$

Figure 5: Response of Markups to Demand Shocks in the 1st and 10th deciles of the markup distribution.

(a) MPS (b) EBP Shocks



Notes: This figure shows the impulse responses of firm-level markups to identified aggregate demand shocks, separately for firms in the 1st (D1) and 10th (D10) deciles of the permanent markup distribution. Panel (a) shows responses to a 1% monetary policy surprise (MPS); Panel (b) shows responses to an EBP shock which raises Chile's CEMBI by 100 b.p. Responses are estimated over the 2010–2019 period, excluding the Great Recession and COVID-19.

Figure 5 shows the cumulative markup response by decile to demand shocks. Panel 5a displays the response to MPS, while Panel 5b shows the response to EBP shocks. In both cases, the markup response is highly asymmetric: firms in the highest decile (D10) exhibit strong and persistent increases in markups, while those in the lowest decile (D1) respond weakly or not. After 24 months, the cumulative markup change for D10 firms is approximately five times larger than that for D1 firms. This pattern confirms that high-markup firms with greater market power and likely more price rigidity drive most of the observed countercyclical response in aggregate markups.

Discussion. In this section, we find robust evidence of the countercyclicality of markups at the firm level. This result has implications for New Keynesian models that rely on this result and the existence of a New Keynesian Phillips curve. Different from other studies, as we mentioned above, we can study all types of firms (not just large-listed firms as in Chiavari et al., 2021), all sectors (unlike Anderson et al., 2018 who study this question with retail data) and at a monthly frequency (unlike Chiavari et al., 2021 with quarterly data and Hong, 2017 with yearly data). A second finding is that firms with greater market power have more countercyclical markups and, thus, contribute more to the aggregate markup cyclicality. These findings suggest that heterogeneous price rigidity and/or firm-level market power are key in transmitting aggregate demand shocks.

This last fact motivates the theoretical framework in the next section, where we model how firms with different levels of market power transmit demand shocks to aggregate outcomes depending on their markup responses.

4 A New Keynesian Model with Heterogeneous Markups

We present a general equilibrium model with endogenous firm-level market power and price adjustment costs. We rely on Edmond et al. (2023) general equilibrium framework and add Rotemberg (1982) price adjustment costs. The price adjustment cost and wedge gaps determine firm-level NKPC and how spread the firm-level markup distribution shapes the response of aggregate variables.[9]

In the model, a representative household has preferences over consumption and labor and owns all the firms in the economy. A perfectly competitive final good producer aggregates intermediate goods from a continuum of heterogenous firms that engage in imperfect competition. These firms produce differentiated varieties combining labor and the final good, which they use as material inputs. Time is discrete.

[9] Recently, Champion et al. (2023) use a similar framework to study at what extent variable markups are related to higher inflation.

4.1 Model Setup

Households. There is a representative household who has preferences over consumption and labor and owns all the firms in the economy. This household solves

$$\max_{\{C_t,L_t\}_{t=0}^{\infty}} \mathbb{E}_0 \left\{ \sum_{t=0}^{\infty} \left(\prod_{s=0}^{t} \beta_s \right) \left(\log(C_t) - \psi \frac{L_t^{1+\nu}}{1+\nu} \right) \right\}$$

s.t.

$$P_t C_t + B_{t+1} = W_t L_t + R_t B_t + DIV_t,$$

where C_t is real consumption of the final good which is the numeraire. L_t denotes labor, W_t is the nominal wage, and D_t is aggregate nominal profits. B_t is a risk-free bond in zero net supply with R_t its nominal gross rate of return. We follow Fernández-Villaverde et al. (2015) and assume the discount factor is time-varying and follows a law of motion given by

$$\beta_t = \beta^{1-\rho_b} \beta_{t-1}^{\rho_b} \exp(\sigma_b \epsilon_{bt}).$$

We call the shock to the discount factor ϵ_{bt}, a *demand shock*. This shock, as we will see next, works as a shock to the Euler equation which has real effects as long as there are nominal rigidities.

The first order conditions of this optimization problem are given by the Euler equation

$$\lambda_t = \mathbb{E}_t \beta_{t+1} \lambda_{t+1} \frac{R_t}{\Pi_{t+1}}$$

with $\lambda_t = \frac{1}{C_t}$ the lagrange multiplier associated to time-t budget constraint and the labor supply given by

$$\psi L_t^{\nu} = \lambda_t \frac{W_t}{P_t}.$$

Final-good producers and market structure. The final good is produced by a perfectly competitive firm using a bundle of unit mass intermediate goods indexed by i. We assume the aggregation of these intermediate goods is done with a Kimball (1995) aggregator, given by

$$\int_0^1 \Upsilon\left(\frac{y_{it}}{Y_t}\right) di = 1, \tag{12}$$

where the function $\Upsilon\left(\frac{y_{it}}{Y_t}\right)$ is strictly increasing and strictly concave. Each period, the representative firm chooses its demand for intermediate goods to maximize profits

$$\max_{\{y_{it}\}_{i=0}^{1}} P_t Y_t - \int_0^1 p_{it} y_{it} di,$$

subject to Equation (12). The solution of this problem gives rise to a demand for each intermediate good

$$\frac{p_{it}}{P_t} = \Upsilon'\left(q_{it}\right) D_t^{-1}, \tag{13}$$

with $D_t = \int \Upsilon'\left(q_{it}\right) q_{it} di$ a demand index where we denoted by $q_{it} = \frac{y_{it}}{Y_t}$ the market share each firm i has in period t. Thus, the demand elasticity is given by

$$\sigma(q_{it}) = \frac{\Upsilon'\left(q_{it}\right)}{\Upsilon''\left(q_{it}\right) q_{it}} \tag{14}$$

A functional form for $\Upsilon(q_{it})$. As in Edmond et al. (2023), we follow Klenow and Willis (2016) and assume inverse demand curves given by:

$$\Upsilon'(q) = \frac{\bar{\sigma} - 1}{\bar{\sigma}} \exp\left(\frac{1 - q^{\varepsilon/\bar{\sigma}}}{\varepsilon}\right), \quad \bar{\sigma} > 1,$$

that imply that the demand elasticity $\sigma(q)$ is log-linear in relative size:

$$\sigma(q) = -\frac{\Upsilon'(q)}{\Upsilon''(q)q} = \bar{\sigma} q^{-\varepsilon/\bar{\sigma}}. \tag{15}$$

As suggested by Edmond et al. (2023), $\varepsilon/\bar{\sigma}$ is the sensitivity of the demand elasticity with respect to relative size of firms, an object they dub the *superelasticity*. As ε increases, large firms face relatively more inelastic demands than smaller firms and thus are able to charge larger markups. While if $\varepsilon = 0$ all firms face a constant demand elasticity and charge the same markup.

Markups with flexible prices. In the flexible price equilibrium, after maximization, firms with monopolistic power set prices given by a *desired* markup over their marginal cost,

$$p_{it}^{flex} = \mu_{it}^{d} mc_{it}.$$

With Kimball aggregator, the flexible price markup depends on the market share; specifically, $\mu_{it}^{d} = \mu^{d}(q_{it})$. Combining this idea with the expression for the demand elasticity, Equation (15), we can express the desired markup as

$$\mu_{it}^{d} = \frac{\sigma_{it}}{\sigma_{it} - 1} = \frac{\overline{\sigma}}{\overline{\sigma} - q_{it}^{\varepsilon/\sigma}}, \tag{16}$$

which shows that markups in the flexible prices equilibrium depend negatively on market shares q_{it} whenever $\varepsilon/\overline{\sigma} > 0$. Therefore, our flexible price equilibrium also features time-varying markups at the firm level.

Intermediate goods producers: Technology. Intermediate firms produce each variety i and are heterogeneous in their productivity z_i. They are engaged in imperfect competition, and produce using labor l_{it} with a constant returns to scale production function:

$$y_{it} = z_i l_{it}.$$

The total cost function is given by:

$$C_{it}(y_{it}) = z_i^{-1} W_t y_{it}.$$

And finally, the implied marginal cost is:

$$mc_{it} = \frac{W_t}{z_i}.$$

Firm productivity z_i is drawn from a Pareto distribution with shape parameter $\kappa > 1$ and lower bound $z_{\min} > 0$:

$$F(z) = 1 - \left(\frac{z_{\min}}{z}\right)^{\kappa}, \quad \text{for } z \geq z_{\min}.$$

This distribution captures heterogeneity in production efficiency across firms and gives rise to a fat-tailed distribution of revenue and firm size. Assuming Pareto-distributed productivities implies that marginal costs and firm sizes are highly skewed, with a small number of highly productive firms accounting for a disproportionate share of output. The Pareto structure ensures tractability of aggregates under heterogeneity and yields closed-form expressions for price indices and general equilibrium allocations.

Intermediate goods producers: Price setting. Each intermediate producer i chooses its price to maximize profits subject to Rotemberg (1982) price adjustment costs. These costs are quadratic on price changes and expressed as a function of sales $p_{it}y_{it}$; this is, $\frac{\theta}{2}\left(\frac{p_{it}}{p_{it-1}}-1\right)^2 p_{it}y_{it}$, where θ is the parameter that drives the degree of price rigidity, which is common to all firms. Therefore, each intermediate producer i chooses prices to solve

$$\max_{\{p_{it}\}_{t\geq 0}} \mathbb{E}_0 \sum_{t=0}^{\infty}\prod_{s=0}^{t}\left(\frac{1}{1+r_s}\right)\left\{(p_{it}-mc_{it})\,y_{it}-\frac{\theta}{2}\left(\frac{p_{it}}{p_{it-1}}-1\right)^2 p_{it}y_{it}\right\}.$$

Given the assumptions above, firm i price changes (after the intermediate firms optimization) are determined by the following firm-level New Keynesian Phillips curve:

$$dp_{it}(dp_{it}-1)=\frac{\sigma_{it}}{\theta}\left(\frac{1}{\mu_{it}}-\frac{1}{\mu_{it}^{d}}\right)+\left(\frac{1}{1+r_{t+1}}\right)dp_{it+1}(dp_{it+1}-1). \tag{17}$$

Equation (17) is the firm-level New Keynesian Phillips Curve. Similar to Equation (3) in Section 2 this equation relates firm-level price changes ($dp_{it}=\log(p_{it}/p_{it-1})$) with firm-level markups. We express the firm-level Phillips curve in terms of markup deviations, with $\mu_{it}=p_{it}/mc_{it}$ the observed—or actual— markup and μ_{it}^{d} the desired markup—the markup level that would prevail without price rigidities.

Additionally, the slope of this firm-level NKPC depends on the demand elasticity, which can also be time-varying but, most importantly, heterogeneous. Recall that using the Lerner index, $\sigma_i=\frac{\mu_i}{\mu_i-1}$, we have that firms with higher markups would face a lower demand elasticity, and thus these firms will face a lower slope in their New Keynesian Phillips curve. Another result derived from that fact is that firms that face more rigid prices have more fluctuating markups in response to demand shocks since the left-hand side of the firm-level NKPC does not respond strongly due to a low σ_{it}.[10]

This implies a relationship between pricing behavior from strategic complementari-

[10]See Appendix C for a derivation of Equation (17).

ties, price rigidities, and markup fluctuations. These facts rationalize the stronger markup response to demand shocks for firms with higher markups.

Aggregate markup and aggregate profits. The aggregate markup $\mathcal{M}_t$ is given by a sales-weighted harmonic average of firm-level markups

$$\mathcal{M}_t = \left(\int_0^1 \frac{\omega_{it}}{\mu_{it}} di \right)^{-1},$$

with $\omega_{it} = \frac{p_{it} y_{it}}{P_t Y_t} \sim \Upsilon'(q_{it}) q_{it}$. In equilibrium the endogenous cross-section of sales shares ω_{it} are determined by the exogenous distribution of productivities. With this expression, we can compute aggregate profits

$$\begin{aligned} D_t &= \int_0^1 \left[\left(p_{it} - \frac{\Theta_t}{z_i} \right) y_{it} di - \frac{\theta}{2} \left(\frac{p_{it}}{p_{it-1}} - 1 \right)^2 p_{it} y_{it} \right] di, \\ &= \left(1 - \frac{1}{\mathcal{M}_t} \right) P_t Y_t - \Psi_t, \end{aligned}$$

with $\Psi_t = \int_0^1 \frac{\theta}{2} \left(\frac{p_{it}}{p_{it-1}} - 1 \right)^2 p_{it} y_{it} di$ is aggregate price adjustment costs.

Aggregation and monetary policy. The goods market clearing condition is given by

$$Y_t = C_t + \Psi_t, \tag{18}$$

which is that aggregate production equals aggregate consumption plus aggregate adjustment costs.

Finally, monetary policy follows a Taylor rule that targets inflation

$$R_t = \frac{1}{\beta} \Pi_t^{\phi_\pi} \exp(\epsilon_t^{mp}),$$

where ϵ_t^{mp} is an exogenous AR(1) monetary policy shock.

5 Model Calibration and Quantification

In this section, we first describe how we calibrate the model using Chilean data, and then we show how markup heterogeneity affects the transmission of demand shocks. We show that persistence of shocks is key to obtaining significant amplification effects from Kimball aggregators in this setting.

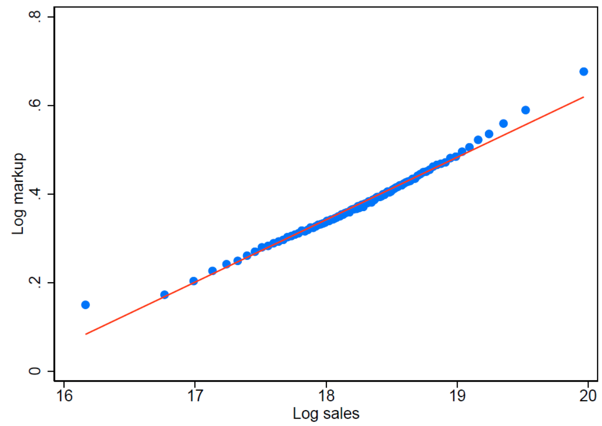
5.1 Calibration

To maximize the model's match to the data, we calibrate three parameters using the Chilean data and obtain the remaining from the literature. In particular, we calibrate the aggregate markup, the super elasticity, and the Pareto tail parameter for the firm-level productivity distribution.

Aggregate Markup. With Kimball demand, where markups arise endogenously, the harmonic mean of markups weighted by revenue shares accurately captures the welfare-relevant distortions, as it reflects how price-cost gaps translate into real output efficiency. Unlike the arithmetic mean, which treats all markups equally and overstates the influence of high-markup firms—especially when markups increase with firm size—the harmonic mean gives more weight to firms pricing closer to marginal cost and thus contributing more efficiently to aggregate output. What matters for aggregate outcomes is not the average markup itself, but how much each firm's markup distorts its marginal cost relative to its share in total revenue. We compute the harmonic mean markup for every month-year for the 2005-2022 period and take the mean accros time to get an aggregate markup of 1.29.

Superelasticity. We follow Edmond et al. (2023) to estimate the empirical relationship between firm-level markups and size, measured by the firm's share of aggregate sales. The slope of this relationship—the "superelasticity"—governs how market power varies with size. To visualize this, we begin with a binscatter plot relating log markups (which we estimate before) to log sales shares, controlling for firm fixed effects. As shown in Figure 6, we observe a clear positive relationship: larger firms charge systematically higher markups.

Figure 6: Log markup vs log sales



Notes: This figure plots 100 bins of the relationship between firm fixed-effect residualized log markups and log sales shares. Firm fixed effects remove persistent differences across firms. The upward slope indicates that larger firms tend to have higher markups.

To estimate the superelasticity, we follow the structural pricing condition in Edmond et al. (2023) and run the following regression:

$$\frac{1}{\mu_{it}} + \log\left(1 - \frac{1}{\mu_{it}}\right) = \alpha + b \log \omega_{it} + \Omega + \nu_{it}, \quad \text{with} \quad b = \frac{\varepsilon}{\bar{\sigma}}. \tag{19}$$

The left-hand side is a nonlinear transformation of the observed markup μ_{it} that arises from firm optimization under Kimball demand. The regression coefficient b identifies the superelasticity: the elasticity of a firm's demand elasticity with respect to its market share. A higher value of b implies stronger markup-size sensitivity, meaning that larger firms face less elastic demand and are able to charge higher markups.

Table 2 presents the results. Across all specifications, the estimated superelasticity is positive and statistically significant. Column (1) includes firm fixed effects only, yielding a baseline estimate of 0.096. As we add time-varying controls in Columns (2) and (3)—firm-by-year and sector-by-month fixed effects—the superelasticity estimate increases to around 0.27. These richer fixed effects control for aggregate shocks and sectoral trends, suggesting that the observed markup-size gradient is not driven by compositional or time-varying factors. The stability of the estimates across specifications reinforces the robustness of the empirical relationship.

Table 2: Superelasticity Estimates

	(1)	(2)	(3)
$\log \omega_{it}$	0.096	0.266	0.269
	(0.001)	(0.002)	(0.002)
Firm FE	✓		
Firm × Year FE		✓	
Firm × Year + Sector × Month FE			✓
R^2	0.32	0.48	0.49
N observations	3,155,394	3,082,571	3,079,592

Notes: The dependent variable is the transformed markup term derived from the firm's first-order condition under Kimball demand. $\log \omega_{it}$ is the firm's log sales share in month t. Standard errors are clustered at the firm level. The coefficient b identifies the superelasticity $\varepsilon/\bar{\sigma}$, which governs how markups rise with firm size in the model.

These estimates imply a non-negligible degree of heterogeneity in demand elasticities across firms. In the model, this heterogeneity gives rise to an endogenous distribution of markups, flatter firm-level Phillips curves for large firms, and uneven responses to aggregate shocks.

Productivity Pareto Tail. To discipline the distribution of firm productivity, we estimate the shape parameter κ of a Pareto distribution using firm-level revenue data and an estimate of the super-elasticity of demand. Under Kimball demand, revenue maps to productivity via a power law determined by the curvature of the demand function. We estimate the Pareto tail of revenue using a maximum likelihood estimator and recover κ by adjusting for demand curvature. Appendix D provides the details of the estimation procedure, where we find that the Pareto tail parameter is 3.93.

Remaining Parameters. The remaining parameters are presented in Table 3, where we set the discount factor to generate an annual interest rate of 4% in steady state, the Frisch elasticity of labor demand is set to 1, the same value that the disutility of labor ψ takes. The Taylor rule coefficient is set to $\phi_\pi = 1.5$, a standard value in the literature. Finally, we set the (common) parameter of price adjustment costs to $\theta = 250$, to deliver a relative response of inflation to output of half in the baseline calibration.

Table 3: Model Parameters

	Value	Target/source
Time discounting	$\beta = 0.99$	Standard
Taylor rule coeff.	$\phi_\pi = 1.5$	Standard
Frisch Elasticity	$\frac{1}{\nu} = 1$	Standard
Disutility of labor	$\psi = 1$	Standard
Elasticity of substitution	$\overline{\sigma} = 5.3$	Match Agg. markup 1.29
Super elasticity	$\varepsilon = 0.27$	See Table 2
Productivity distribution Pareto tail	$\kappa = 3.93$	See Appendix D
Price Rigidity	$\theta = 250$	Match rel. response of Y_t and π_t

5.2 Dynamic Responses to Demand Shocks

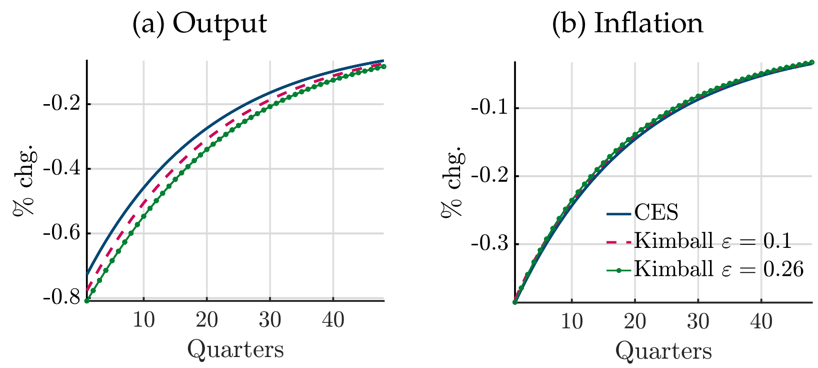
This subsection quantifies the role of demand-side complementarities in amplifying the macroeconomic effects of contractionary demand shocks. We show that even modest levels of superelasticity in demand substantially increase the output response to shocks, with minimal effect on inflation. The key amplification mechanism operates through endogenous and persistent dispersion in firm-level markups: higher superelasticity leads to greater heterogeneity in price adjustments, especially among high-markup firms. This dispersion is not only amplified by general equilibrium forces but also becomes state-dependent—larger and more persistent when the economy is hit by serially correlated shocks. Together, these results underscore that heterogeneity in pricing power plays a critical role in shaping aggregate dynamics in response to demand shocks.

Figure 7 depicts the impulse responses of output and inflation to a contractionary demand shock. We compare three calibrations to evaluate the role of the superelasticity of demand in shaping the aggregate responses: (i) a baseline calibration with $\varepsilon = 0.27$; (ii) a lower superelasticity of $\varepsilon = 0.1$; and (iii) a CES benchmark, which approximates the limiting case $\varepsilon \to 0$. The shock is modeled as a highly persistent discount factor shock, with persistence parameter $\rho_b = 0.95$.

In all cases, the shock leads to a decline in both inflation and output. However, the magnitude of the output response is amplified when the superelasticity of demand is higher. In particular, with $\varepsilon = 0.27$, the initial decline in output is 11% larger compared to the CES benchmark, and this effect remains persistent over time. We also consider an intermediate calibration with $\varepsilon = 0.1$, which better approximates estimates for the U.S.

economy (see Edmond et al., 2023), and find that even modest levels of strategic complementarities in demand amplify the response to demand shocks. These results are noteworthy because they emerge in an environment with perfect labor mobility across firms. In settings with input frictions or imperfect mobility—where reallocation from low- to high-markup firms is more limited—the amplification effects of demand-side complementarities would likely be even stronger.

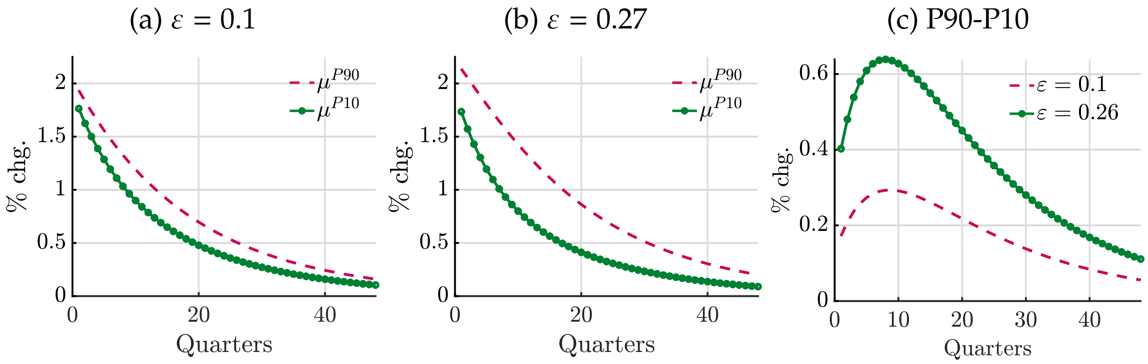
Figure 7: Responses output and inflation for different superelasticities.



Notes: This figure shows the responses of output (Panel a) and inflation (Panel b) to a contractionary demand shock under three alternative values of the superelasticity of demand ε: (i) baseline calibration with $\varepsilon = 0.27$; (ii) lower superelasticity $\varepsilon = 0.1$; and (iii) a CES benchmark approximating $\varepsilon \to 0$. The demand shock is modeled as a highly persistent discount factor shock with persistence $\rho_b = 0.95$.

Our amplification results operate through endogenous markup dispersion. Figure 8 reports the responses of markups at the 10th and 90th percentiles of the firm-level markup distribution. Panel 8b shows that when the superelasticity is high, markups for all firms rise by a larger amount—primarily through general-equilibrium effects—with especially pronounced and persistent increases in the upper tail (P90) relative to the low-superelasticity case in Panel 8a. Consequently, the gap between high- and low-markup firms widens and remains elevated: Panel 8c documents a sizeable and persistent increase in markup dispersion in the high-superelasticity calibration. This endogenous, persistent dispersion is the key channel through which strong strategic complementarities amplify aggregate responses to demand shocks.

Figure 8: Responses of P10 and P90 markups and dispersion for different superelasticities.

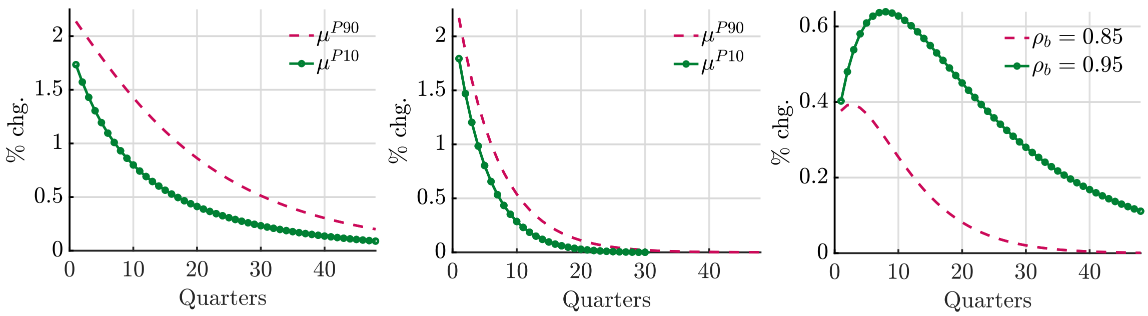


Notes: This figure displays the response of firm-level markups at the 10th and 90th percentiles of the distribution, as well as the dispertion (P90–P10), following a contractionary demand shock under two superelasticity calibrations. Panel (a) shows the low-superelasticity case ($\varepsilon = 0.1$), while Panel (b) shows the high-superelasticity case ($\varepsilon = 0.27$). Panel (c) reports the resulting markup dispersion (P90–P10).

The mechanism is state-dependent. After an initial shock increases dispersion, subsequent shocks hit an economy whose markup distribution is already more spread out, compounding the aggregate effects. Figure 9 illustrates that this compounding is more pronounced when shocks are highly persistent; with lower shock persistence, the rise in dispersion is both milder and shorter-lived.

Figure 9: Responses of P10 and P90 markups and dispersion for different shock's persistence.

(a) ρ_b = 0.95 (b) ρ_b = 0.85 (c) P90-P10



Notes: This figure examines how the persistence of a contractionary demand shock affects markup dispersion. Panels (a) and (b) show the response of P10 and P90 firm-level markups for two values of shock persistence: ρ_b = 0.95 (high) and ρ_b = 0.85 (low). Panel (c) shows the corresponding markup dispersion (P90–P10). When shocks are more persistent, markup dispersion rises more sharply and remains elevated for longer. This illustrates the compounding mechanism: persistent shocks interact with endogenous markup dispersion, generating stronger amplification over time.

Our quantification exercises highlight that endogenous markup dispersion is a powerful amplification channel for aggregate demand shocks. When firms differ in their pricing power, and those with higher markups adjust prices less, contractionary shocks widen the distribution of markups and depress output more strongly—especially in the presence of strategic complementarities in demand. This amplification is further intensified when shocks are persistent, as the economy accumulates dispersion over time. These findings reinforce the empirical evidence and theoretical mechanisms presented earlier: heterogeneity in firm markups is not a second-order detail but a core driver of macroeconomic dynamics. Accounting for these features is essential for understanding the asymmetric transmission of shocks and for designing effective stabilization policies.

6 Conclusion

This paper shows that firm-level heterogeneity in markups amplifies the effects of aggregate demand shocks. We derive a New Keynesian Phillips Curve that highlights how inflation dynamics are shaped by the dispersion in markups, not just their average. Firms with higher markups adjust prices less, leading to a more pronounced and persistent increase in markup dispersion following a negative demand shock.

Using Chilean administrative data, we provide empirical support for this mechanism: markups are strongly countercyclical, particularly among high-markup firms, and dispersion rises during downturns. We embed these features in a general equilibrium model and show that strategic complementarities in demand, captured by a positive superelasticity, amplify output responses by up to 20% relative to a CES benchmark. This amplification is stronger when shocks are more persistent.

Our results have clear policy implications. First, monetary policy may be more effective in stabilizing output than inflation in environments with high markup dispersion, as price rigidity becomes increasingly asymmetric. Second, since the effects of demand shocks are amplified when factor reallocation is limited, policies that enhance input mobility, such as reducing labor market frictions or improving access to capital, can reduce macroeconomic volatility. Finally, promoting competition may help reduce the persistence of markup dispersion, enhancing the transmission of stabilization policy.

Our results open at least three avenues for future work. One direction is to explore how input frictions, such as labor immobility or credit constraints, interact with markup dispersion to further amplify aggregate fluctuations. A second one is to investigate the role of endogenous firm entry and exit in shaping the dynamics of markup heterogeneity over the business cycle. Finally, empirical work in other countries with detailed firm-level data could assess the external validity of these findings and evaluate how institutional environments affect the strength of these amplification mechanisms.

Appendix

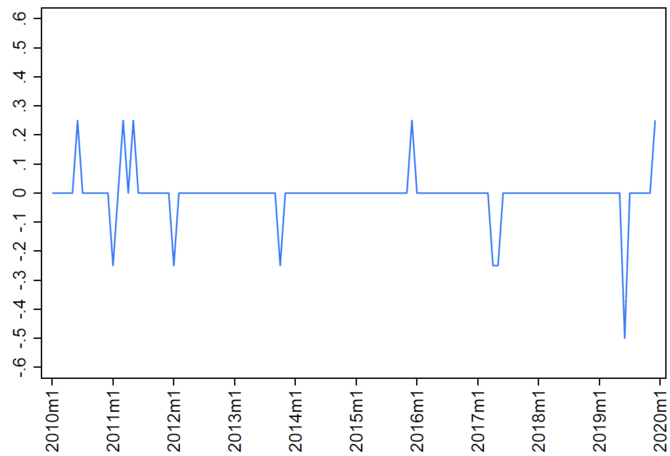
A Additional Empirical Facts

Table 4: Main production variables by industry (Median)

	Revenue	Material expenditure	Wage bill	# employees
Agriculture	45	21	10	16
Mining	92	56	15	19
Manufacturing	66	38	15	21
Utilities	195	96	37	32
Construction	45	20	12	16
Retail and Wholesale	77	50	13	19
Transport and ICTs	65	25	18	20
Financial Services	127	32	57	22
Real Estate Services	58	17	21	17
Business Services	56	13	23	22
Personal Services	43	13	23	22

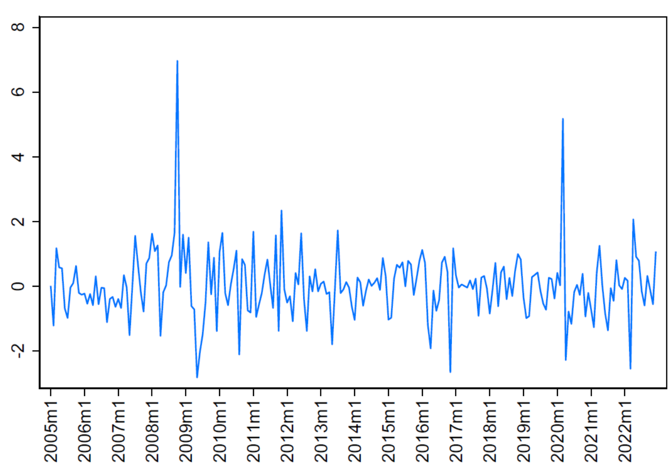
Notes: This table reports the median values of main production variables—revenue, material expenditure, wage bill, and number of employees—by industry in 2018, using firm-level data from Chile. All monetary values are in millions of nominal 2018 Chilean pesos; the number of employees is a headcount. Utilities and Financial Services exhibit the highest median revenues (195M and 127M, respectively), reflecting their capital intensity and large transaction volumes. Utilities also lead in material expenditure and employee count, while Financial Services show the highest median wage bill (57M), consistent with their high labor compensation. Manufacturing and Mining have relatively high material intensity (materials account for more than half of revenue in Mining). Business and Personal Services have the lowest material expenditures but relatively high wage bills per worker, indicating labor-intensive production. These differences underscore substantial heterogeneity in input intensity and scale across industries.

Figure 10: Monetary Surprises mps_t



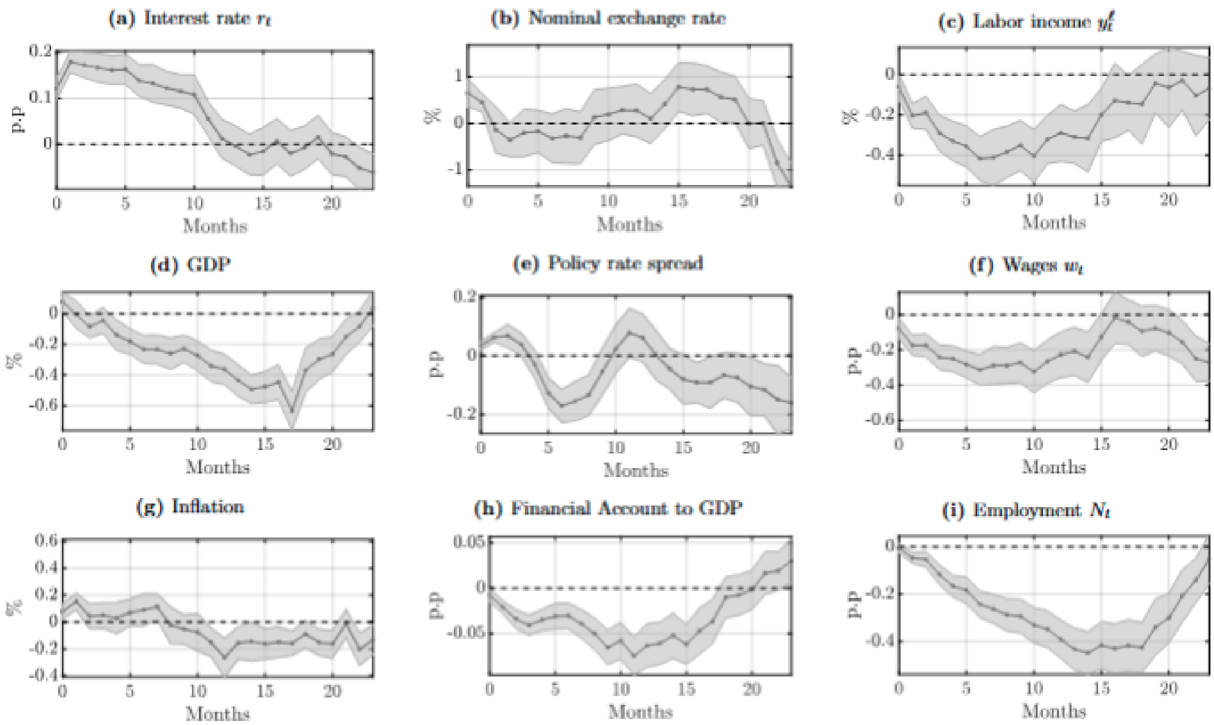
Notes: This figure shows the time series of monetary policy surprises mps_t, constructed as in Aruoba et al. (2021), using the difference between the Central Bank of Chile's policy rate decision and the expected rate from Bloomberg analyst surveys on the same day. Positive values indicate contractionary surprises (tightening beyond expectations), while negative values indicate expansionary surprises.

Figure 11: EBP shocks



Notes: This figure displays the time series of Excess Bond Premium (EBP) shocks used as aggregate financial demand shocks in our empirical analysis. The shocks are constructed following Aldunate et al. (2025), who estimate a monthly structural VAR for the U.S. economy and isolate innovations to the EBP component from Gilchrist and Zakrajšek (2012), controlling for other macroeconomic and financial variables. These shocks are then used as external instruments affecting Chilean financial conditions via global capital markets. Positive values correspond to unexpected increases in risk premia, which tighten financial conditions and reduce aggregate demand.

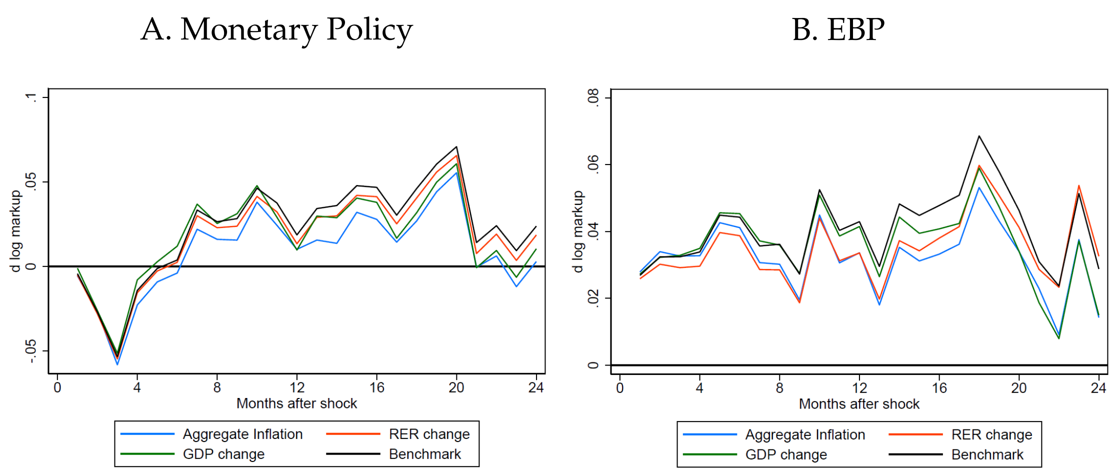
Figure 12: Response of Macroeconomic Variables to EBP shocks



Notes: This figure shows the impulse responses of Chilean macroeconomic variables to a one standard deviation Excess Bond Premium (EBP) shock, based on the identification strategy of Aldunate et al. (2025). The EBP shock is constructed from a structural VAR for the U.S. and captures unexpected increases in credit spreads not driven by fundamentals. A positive EBP shock leads to a contraction in GDP and inflation, consistent with a decline in aggregate demand. The exchange rate depreciates and corporate credit spreads widen, indicating tighter financial conditions.

B Local Projections Robustness

Figure 13: Average Effect Robustness



Notes: This figure presents robustness checks for the average effect of identified aggregate demand shocks—monetary policy surprises (Panel A) and Excess Bond Premium (EBP) shocks (Panel B)—on firm-level markup dynamics. Each line plots the estimated coefficient β_h from Equation **??**, capturing the average cumulative change in log markups h months after the shock. The regressions control for firm-level employment and include firm-month fixed effects; standard errors are clustered at the firm level. The results confirm that the countercyclical markup response is robust to the inclusion of additional macroeconomic controls such as GDP growth, inflation, and the real exchange rate.

C Firm-level NKPC derivation

Firms maximize their profits, subject to a demand with a price elasticity given by σ_{it} and are sibject to Rotemberg (1982) price adjustment costs. They, as in most of the literature on price rigidities with Rotemberg, do not internalize the effect of higher sales on the adjustment costs.

$$\sum_{t=0}^{\infty} \prod_{s=0}^{t} \left(\frac{1}{1+r_s}\right) \left\{ (p_{it} - mc_{it})\, y_{it} - \frac{\theta}{2} \left(\frac{p_{it}}{p_{it-1}} - 1\right)^2 p_{it} y_{it} \right\}$$

FOC

$$y_{it} + (p_{it} - mc_{it}) \frac{\partial y_{it}}{\partial p_{it}} - \theta \left(\frac{p_{it}}{p_{it-1}} - 1\right) \frac{p_{it}}{p_{it-1}} y_{it} + \mathbb{E}_t \frac{1}{1+r_{t+1}} \theta \left(\frac{p_{it+1}}{p_{it}} - 1\right) \frac{p_{it+1}^2}{p_{it}^2} y_{it+1} = 0$$

$$\left(\frac{p_{it}}{p_{it-1}}-1\right)\frac{p_{it}}{p_{it-1}}=\frac{1}{\theta}\left(1+\left(1-\frac{mc_{it}}{p_{it}}\right)\frac{\partial y_{it}}{\partial p_{it}}\frac{p_{it}}{y_{it}}\right)+\mathbb{E}_t\frac{1}{1+r_{t+1}}\left(\frac{p_{it+1}}{p_{it}}-1\right)\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$$

$$dp_{it}\left(dp_{it}-1\right)=\frac{1}{\theta}\left(1+\left(1-\frac{mc_{it}}{p_{it}}\right)\frac{\partial y_{it}}{\partial p_{it}}\frac{p_{it}}{y_{it}}\right)+\mathbb{E}_t\frac{1}{1+r_{t+1}}\left(dp_{it+1}-1\right)dp_{it+1}\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$$

Then, using the definition of the inverse of price elasticity $\sigma_{it}=-\frac{\partial y_{it}}{\partial p_{it}}\frac{p_{it}}{y_{it}}$, we have

$$dp_{it}(dp_{it}-1)=\frac{1}{\theta}\left(1-\left(1-\frac{mc_{it}}{p_{it}}\right)\sigma_{it}\right)+\mathbb{E}_t\frac{1}{1+r_{t+1}}dp_{it+1}(dp_{it+1}-1)\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$$

And

$$dp_{it}(dp_{it}-1)=\frac{\sigma_{it}}{\theta}\left(\frac{mc_{it}}{p_{it}}-\frac{\sigma_{it}-1}{\sigma_{it}}\right)+\mathbb{E}_t\frac{1}{1+r_{t+1}}dp_{it+1}(dp_{it+1}-1)\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}},$$

Loglinearization. With $\frac{1}{\mathcal{M}_{it}}=\frac{mc_{it}}{p_{it}}$ and rearranging

$$\ln(\sigma_{it}-1)=\ln\left(\sigma_{it}\frac{1}{\mathcal{M}_{it}}-\theta dp_{it}(dp_{it}-1)+\theta\mathbb{E}_t\frac{1}{1+r_{t+1}}dp_{it+1}(dp_{it+1}-1)\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}\right)$$

$$\frac{d\sigma_{it}}{\sigma_i-1}=\frac{1}{\sigma_i-1}\Bigg[d\sigma_{it}\frac{1}{\mathcal{M}_i}-\sigma_i\frac{d\mathcal{M}_{it}}{\mathcal{M}_i^2}-\theta(dp_id(dp_{it})+d(dp_{it})(dp_i-1))+\theta\beta\mathbb{E}_tdp_id(dp_{it+1})$$
$$+\theta\beta\mathbb{E}_tdp_idp_{it+1}+\theta\beta\mathbb{E}_tdp_i(dp_i-1)d\left(\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}\right)\Bigg]$$

With $dp_i=1$:

$$\frac{d\sigma_{it}}{\sigma_i-1}=\frac{1}{\sigma_i-1}\left[d\sigma_{it}\frac{1}{\mathcal{M}_i}-\sigma_i\frac{d\mathcal{M}_{it}}{\mathcal{M}_i^2}-\theta d(dp_{it})+\theta\beta\mathbb{E}_td(dp_{it+1})\right]$$

Then, using the definition of markups, $\mathcal{M}_i=\frac{\sigma_i}{\sigma_i-1}$ several times, we have

$$dp_{it} = \frac{\sigma_i - 1}{\theta}\left(\frac{d\mathcal{M}_{it}^d}{\mathcal{M}_i} - \frac{d\mathcal{M}_{it}}{\mathcal{M}_i}\right) + \beta\mathbb{E}_t dp_{it+1}$$

$$dp_{it} = -\frac{\sigma_i - 1}{\theta}\tilde{\mu}_{it} + \beta\mathbb{E}_t dp_{it+1}$$

with $\tilde{\mu}_{it} = \left(\frac{d\mathcal{M}_{it}^d}{\mathcal{M}_i} - \frac{d\mathcal{M}_{it}}{\mathcal{M}_i}\right)$

D Pareto Tail of Productivity Distribution Estimation

We aim to estimate the shape parameter κ of the Pareto distribution for firm productivity z_i, under the assumption that firms face Kimball demand. We observe firm-level revenue R_i, have estimated the super-elasticity parameter $\varepsilon/\bar{\sigma}$ from the markup distribution, assume that firms are monopolistic competitors with labor-only production, and that the revenue function is isoelastic: $R_i = Ay_i^\gamma$, where $0 < \gamma < 1$.

We proceed in four steps:

Step 1: Estimate the Pareto tail of firm revenue. Assuming that firm revenue follows a Pareto distribution in the upper tail,

$$\mathbb{P}(R_i > r) \sim r^{-\alpha}, \quad \text{for } R_i \geq R_{\min},$$

the maximum likelihood estimator for the tail index α is:

$$\hat{\alpha} = \left[\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{R_i}{R_{\min}}\right)\right]^{-1},$$

where $R_{\min}$ is a threshold defining the lower bound of the Pareto tail, and the sum is taken over firms with $R_i \geq R_{\min}$.

Step 2: Relate revenue to productivity using demand curvature. Under Kimball demand, with super-elasticity $\varepsilon/\bar{\sigma}$, the revenue function is:

$$R_i = Ay_i^{\gamma}, \quad \text{where } \gamma = \frac{\sigma - 1}{\sigma - 1 + \varepsilon/\bar{\sigma}},$$

and σ is the elasticity of substitution in the CES benchmark.

Combining this with the production function $y_i = z_i l_i$ and cost minimization, we obtain the mapping:

$$z_i \propto R_i^{\xi}, \quad \text{where } \xi = \frac{1-\gamma}{\gamma}.$$

Step 3: Infer the Pareto tail parameter of productivity. Since $z_i \propto R_i^{\xi}$, and R_i is Pareto distributed with shape α, then z_i follows a Pareto distribution with shape:

$$\kappa = \frac{\alpha}{\xi} = \alpha \cdot \frac{\gamma}{1-\gamma}.$$

Substituting the expression for γ in terms of $\varepsilon/\bar{\sigma}$, we obtain:

$$\kappa = \hat{\alpha} \cdot \frac{\sigma - 1}{\varepsilon/\bar{\sigma}}.$$

Step 4: Final estimator. Given observed R_i, the estimated super-elasticity $\varepsilon/\bar{\sigma}$, and a chosen threshold $R_{\min}$, the final estimator is:

$$\hat{\kappa} = \left[\frac{1}{N}\sum_{i=1}^{N} \log\left(\frac{R_i}{R_{\min}}\right)\right]^{-1} \cdot \frac{\sigma - 1}{\varepsilon/\bar{\sigma}}, \quad \text{for } R_i \geq R_{\min}.$$

This provides a structural estimate of the Pareto tail parameter of firm productivity using revenue data and the estimated super-elasticity of demand.

References

ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): "Identification properties of recent production function estimators," *Econometrica*, 83, 2411–2451.

AFROUZI, H. AND L. CALOI (2023): "Endogenous firm competition and the cyclicality of markups," *Review of Economics and Statistics*, 1–45.

ALDUNATE, R., A. BLANCO, A. FERNANDEZ, M. GIARDA, AND G. NAVARRO (2025): "Cross-Sectional Labor Dynamics After a Foreign Shock," *Mimeo, Central Bank of Chile.*

ANDERSON, E., S. REBELO, AND A. WONG (2018): "Markups Across Space and Time," Working Paper 24434, National Bureau of Economic Research.

ARUOBA, B., A. FERNÁNDEZ, D. GUZMÁN, E. PASTÉN, AND F. SAFFIE (2021): "Monetary Policy Surprises in Chile: Measurement & Real Effects," Working Papers Central Bank of Chile 921, Central Bank of Chile.

BAQAEE, D. R., E. FARHI, AND K. SANGANI (2024): "The supply-side effects of monetary policy," *Journal of Political Economy*, 132, 1065–1112.

BILS, M., P. J. KLENOW, AND B. A. MALIN (2018): "Resurrecting the role of the product market wedge in recessions," *American Economic Review*, 108, 1118–1146.

BOND, S., A. HASHEMI, G. KAPLAN, AND P. ZOCH (2021): "Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data," *Journal of Monetary Economics*.

BURSTEIN, A., V. M. CARVALHO, AND B. GRASSI (2020): "Bottom-up Markup Fluctuations," Working Paper 27958, National Bureau of Economic Research.

CHAMPION, M., C. EDMOND, AND J. HAMBUR (2023): "Competition, Markups, and Inflation: Evidence From Australian Firm-Level Data Preliminary Draft," .

CHIAVARI, A., D. SMIRNOV, AND M. MORAZZONI (2021): "Heterogeneous markups cyclicality and monetary policy," Tech. rep., Working paper.

DE LOECKER, J. (2021): "Comment on Bond et al (2021)," *Journal of Monetary Economics*.

DE LOECKER, J. AND F. WARZYNSKI (2012): "Markups and firm-level export status," *American economic review*, 102, 2437–71.

DE RIDDER, M., B. GRASSI, G. MORZENTI, ET AL. (2022): "The hitchhiker's guide to markup estimation," Tech. rep.

DE ROUX, N., M. ESLAVA, S. FRANCO, AND E. VERHOOGEN (2021): "Estimating production functions in differentiated-product industries with quantity information and external instruments," Tech. rep., National Bureau of Economic Research.

DHYNE, E., A. PETRIN, V. SMEETS, AND F. WARZYNSKI (2022): "Theory for extending single-product production function estimation to multi-product settings," Tech. rep., National Bureau of Economic Research.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2023): "How costly are markups?" *Journal of Political Economy*, 131, 1619–1675.

FERNÁNDEZ-VILLAVERDE, J., G. GORDON, P. GUERRÓN-QUINTANA, AND J. F. RUBIO-

RAMÍREZ (2015): "Nonlinear adventures at the zero lower bound," *Journal of Economic Dynamics and Control*, 57, 182–204.

GAGLIARDONE, L., M. GERTLER, S. LENZU, AND J. TIELENS (2023): "Anatomy of the Phillips curve: micro evidence and macro implications," Tech. rep., National Bureau of Economic Research.

GALÍ, J. (2015): *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press.

GALÍ, J., M. GERTLER, AND D. LOPEZ-SALIDO (2007): "Markups, Gaps, and the Welfare Costs of Business Fluctuations," *The Review of Economics and Statistics*, 89, 44–59.

GILCHRIST, S. AND E. ZAKRAJŠEK (2012): "Credit spreads and business cycle fluctuations," *American economic review*, 102, 1692–1720.

HALL, R. E. (1988): "The relation between price and marginal cost in US industry," *Journal of political Economy*, 96, 921–947.

HONG, S. (2017): "Customer Capital, Markup Cyclicality, and Amplification," Working Papers 2017-33, Federal Reserve Bank of St. Louis.

HÖYNCK, C., M. LI, AND D. ZHANG (2023): "The (Unequal) Rise of Market Power, Nominal Rigidities, and Monetary Non-Neutrality," .

KIMBALL, M. S. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 27, 1241–1277.

KLENOW, P. J. AND J. L. WILLIS (2016): "Real rigidities and nominal price changes," *Economica*, 83, 443–472.

KOUVAVAS, O., C. OSBAT, T. REINELT, AND I. VANSTEENKISTE (2021): "Markups and inflation cyclicality in the euro area," .

MEIER, M. AND T. REINELT (2022): "Monetary policy, markup dispersion, and aggregate tfp," *Review of Economics and Statistics*, 1–45.

NEKARDA, C. J. AND V. A. RAMEY (2020): "The cyclical behavior of the price-cost markup," *Journal of Money, Credit and Banking*, 52, 319–353.

OTTONELLO, P. AND T. WINBERRY (2020): "Financial heterogeneity and the investment channel of monetary policy," *Econometrica*, 88, 2473–2502.

ROTEMBERG, J. J. (1982): "Sticky prices in the United States," *Journal of political economy*, 90, 1187–1211.