NBER WORKING PAPER SERIES

INPUT PRICE DISPERSION ACROSS BUYERS AND MISALLOCATION

Ariel Burstein
Javier Cravino
Marco Rojas

Working Paper 33128
http://www.nber.org/papers/w33128

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2024

Input Price Dispersion Across Buyers and Misallocation
Ariel Burstein, Javier Cravino, and Marco Rojas
NBER Working Paper No. 33128
November 2024
JEL No. D22, D24, D4, E02, O4

ABSTRACT

We leverage a comprehensive dataset of electronic invoices from Chilean firms to document new facts on price dispersion across buyers of manufactured intermediate goods. Over half of firm-to-firm manufacturing sales are accounted for by products that are purchased by more than one buyer in the same month. Price dispersion across buyers is pervasive, with a price range across buyers of 46 percentage points for the average product. Price gaps are highly persistent over time and strongly correlated across different products purchased by the same buyer. While price disparities comove with observable characteristics of buyer-seller pairs—such as size of the buyer and the transaction—these factors account for a small portion of the overall variation in price gaps. We use a workhorse model of production networks to quantify the productivity gains from eliminating observed dispersion in prices across buyers of the same product, under the assumption that this dispersion is driven by buyer-product specific markups. The increase in aggregate productivity (normalized by the sales share of treated multi-buyer firms) ranges from 2 to 7 percent, depending on the calibration of elasticities of substitution. The gains from eliminating markup dispersion across buyers are as large as those of eliminating markup dispersion across products.

Ariel Burstein
Department of Economics
Bunche Hall 8365
Box 951477
UCLA
Los Angeles, CA 90095-1477
and NBER
arielb@econ.ucla.edu

Marco Rojas
Research Department
Banco Central de Chile
Agustinas 1180
Santiago
Chile
mrojaso@bcentral.cl

Javier Cravino
Department of Economics
University of Michigan
611 Tappan Street, Lorch Hall 365B
Ann Arbor, MI 48109
and NBER
jcravino@umich.edu

1 Introduction

To what extent do different firms pay disparate prices for the same intermediate input? Which observable characteristics of buyers and sellers account for the variation in prices across buyers? Is dispersion in input prices across buyers an important source of misallocation in domestic production networks? Answering these questions is important for guiding models of firm-to-firm trade and for understanding inefficiencies in domestic supply chains. However, existing evidence on price dispersion for intermediate inputs is scarce and limited to products within narrow industries or to international trade transactions.[1]

This paper uses comprehensive firm-to-firm data from Chile to document new facts on price dispersion across buyers of intermediate inputs and to quantify the extent of misallocation resulting from this price dispersion. Our analysis is based on transaction-level price data extracted from electronic invoices issued by Chilean manufacturing firms between January 2019 and December 2023. Crucially, these invoices include detailed product codes and descriptions that firms use in their billing systems, which are essential to uniquely identify products. They also contain the identities of the sellers and the buyers, which allow us to study how the price of a given product varies across buyers. The set of uniquely identified products represents almost half of Chilean firm-to-firm manufacturing sales. The industry composition and the price changes in our data closely mimic those in the Chilean producer price index.

Manufacturing firms frequently sell the same product to multiple buyers at different prices. Among our sample of products, more than half of firm-to-firm sales are accounted for by products that are purchased by more than one buyer in the same month. For the average product with multiple buyers, the range in log prices across buyers is 0.38 (46 percentage points) and the cross-buyer standard deviation in log prices is 0.08. Dispersion in prices is very stable across months in our sample, which covers a period with substantial variation in the underlying inflation rate.[2] Price dispersion is pervasive across disparate manufacturing products, ranging from Food, to Chemicals, to Electrical Equipment, and is more pronounced for products that are sold by larger sellers, have higher sales volume, and are purchased by a larger number of buyers.

We evaluate how observed 'price gaps', i.e. the difference between the price that an individual buyer pays and the product's average price across buyers, are related to observable characteristics of the buyers, sellers, and products. Price gaps are persistent over

[1] See e.g., Grennan (2013), Marshall (2020), Fontaine et al. (2020), and Alviarez et al. (2023). We discuss this literature below.

[2] Between June 2021 and December 2023, annual PPI inflation in Chile ranged between -5% and 30%.

time, correlated across different products purchased by the same buyer, and across products within buyer-seller pairs. Prices are lower for larger buyers, for buyer-seller pairs with older relations and higher volume of sales, and are decreasing in the volume of purchased quantities. These observable characteristics, however, only account for a small fraction of the variation in price gaps in the data. Whereas sellers charge different prices across buyers, we do not find evidence that buyers face fixed price-quantity menus; buyers often purchase varying quantities while facing constant prices. Finally, we show that while the geographical distance between the buyer and seller is positively correlated to prices, it plays a minimal role in accounting for the observed price gaps. In fact, if we exclude additional surcharges listed on the invoices, the correlation between prices and distance disappears. Moreover, price gaps are not driven by cash payment discounts —if anything, buyers who pay with credit tend to face lower prices. These two last observations suggest that price gaps reflect differences in markups rather than differences in marginal costs of supplying different buyers.

Motivated by these facts, we use a workhorse model of a production network to quantify the welfare costs of markup dispersion across buyers. The model is populated by firms that produce differentiated products using intermediate inputs and a single factor in fixed aggregate supply, and these goods are sold to other firms and final consumers. We introduce markups as exogenous wedges to match observed price gaps in our data. Any model that endogenizes these markups would need to be consistent with the fact that observable characteristics of products and buyers do not account for much of the variation in price gaps. In line with our evidence, we assume that buyers can freely pick quantity at a given price. As in Hsieh and Klenow (2009) and Baqaee and Farhi (2020), dispersion in wedges can distort the allocation of inputs across buyers and reduce allocative efficiency.

We use the model to evaluate changes in aggregate productivity from eliminating markup dispersion across buyers of the same product, while keeping average product-markups unchanged. Under the assumption that the marginal cost of each product does not vary with the identity of the buyer, differences in initial markups across buyers can be measured from observed price gaps. That is, our measure of dispersion in product-buyer wedges does not require estimating markup levels for each product-buyer pair, which requires much stronger assumptions (see e.g., De Loecker, Goldberg, Khandelwal and Pavcnik 2016 and Dhyne, Petrin, Smeets and Warzynski 2022a). Since we do not observe product-level prices for non-manufacturing products and final sales, we keep markups for these products unchanged.

In addition to the change in markups, this counterfactual requires parameterizing the full network of input uses and assigning values to the share of intermediate inputs in production for each product, and the share of each product in total sales. Rather than making distributional assumptions, we target them directly by mapping each product in the model to a product in the Chilean data. We consider alternative values of elasticities of substitution across products and inputs from the literature.

Eliminating markup dispersion across buyers while keeping average product-markups unchanged increases aggregate productivity (and welfare) by 0.12% in the low elasticity of substitution calibration and by 0.37% in the high elasticity of substitution calibration. Since we do not consider markup variation for final sales and for non-manufacturing firms, this counterfactual applies to products that represent only 5% of total sales. The change in aggregate productivity normalized by the fraction of sales of treated firms is 2.3% and 7.1% in the low and high elasticity of substitution calibrations.

To put these gains in perspective, we compare them with a second counterfactual in which we eliminate markup dispersion across both buyers and products. Specifically, starting from setting average product-level markups equal to the observed ratio of revenues to costs, and matching the observed dispersion in prices across buyers, we move to a counterfactual in which markups for all manufacturing intermediate sales are uniformly equal to the economy-wide ratio of revenues to costs. Eliminating markup dispersion across buyers and products results in productivity gains that are roughly double the gains from eliminating markup dispersion across buyers alone. In other words, markup dispersion across buyers is as costly as markup dispersion across products.

Related literature. Most of the evidence on price dispersion is based on consumer prices. For example, Coibion et al. (2015), Kaplan and Menzio (2015), Kaplan et al. (2019), and Sangani (2024) focus on consumer price dispersion within and across retail stores, while DellaVigna and Gentzkow (2019) and Daruich and Kozlowski (2023) document the practice of uniform prices within large retail chains. Evidence of cross-buyer price dispersion for intermediate inputs is largely limited to specific industries or firms. For example, Grennan (2013) studies price dispersion across buyers of medical devices, and Marshall (2020) focuses on wholesale prices paid by different New York restaurants for the same good on the same day. Macedoni and Mattana (2023) quantify dispersion in firm-to-firm prices for life-saving equipment sold by Viking, a Danish multinational.

Other papers examine dispersion of unit values in trade data. Fontaine, Martin and Mejean (2020) use French customs data to document differences in unit values across buyers within CN8 product categories (corresponding to tariff lines used by EU customs) sold by the same exporter. As they note, this variation can partly be explained by product heterogeneity within CN8 categories. In contrast, we use a much narrower definition of products based on product codes and descriptions used in firms' billing systems. If in our dataset we aggregate products at the HS8 tariff lines used by Chilean customs, price dispersion within product-seller across buyers is ten times greater than when using our narrower product definition. Alviarez, Fioretti, Kikkawa and Morlacco (2023) document differences in unit values across US importers purchasing HS10 US customs product categories from the same foreign supplier, while Ignatenko (2023) documents differences in unit values across importers and time in Paraguay. These papers investigate price dispersion using models of oligopoly, oligopsony, and bilateral bargaining. Our empirical contribution relative to these papers is to document price dispersion across buyers in the Chilean firm-to-firm domestic network for the universe of detailed manufactured products, and to show that observed characteristics of buyers, sellers, and transacted quantities have a modest role in accounting for the observed variation of price gaps in our data.

Our paper is also related to an extensive literature studying misallocation due to markup dispersion (see e.g., Peters 2020, Baqaee and Farhi 2020, Edmond et al. 2023, Pellegrino 2023, Osotimehin and Popov 2023).[3] Whereas models in this literature are flexible enough to allow for a buyer-seller specific definition of products, a standard assumption when mapping them to data is that firms charge common markups across all buyers.[4] Closer to us is Dhyne, Kikkawa and Magerman (2022b), who study the gains from eliminating markup dispersion across buyers of the same product, inferring buyer specific markups by combining a model of endogenous markups and market share data. In contrast, we directly infer differences in markups across buyers using data on product-buyer specific prices, under the assumption that marginal cost of production are common to all buyers of the same product. Bornstein and Peter (2023) study the cost of markup dispersion across firms and buyers in a model in which sellers set non-linear prices. Our quantitative model assumes that sellers are able to price discriminate across buyers but, motivated by our data, each buyer faces prices that do not vary with the quantity purchased.

The reminder of the paper is organized as follows. Section 2 introduces the data. Section 3 reports our empirical findings. Section 4 presents our quantitative framework

[3]The welfare cost of inflation in sticky price models depends crucially on how inflation impacts the dispersion of relative prices across goods. Nakamura et al. (2018) and Alvarez et al. (2019) use the micro data underlying the consumer price index to measure the link between inflation and the distribution of price changes. We document very small changes in the dispersion of markups across buyers in Chile as inflation rose sharply in 2022.

[4]Koike-Mori and Martner (2024) quantify the importance of resource misallocation within the firm in measured productivity growth in Chile by extending the non-parametric approach of Baqaee and Farhi (2020) to allow for multi-product firms and joint-production.

and analytical results. Section 5 presents the results of our counterfactual exercises, and the last section concludes.

2 Data

In this section we describe the dataset we use to calculate measures of price dispersion.

2.1 Data description

Our analysis is based on price and quantity data from domestic transactions among firms in Chile. Starting in 2017, the Chilean Internal Tax Service (SII for its acronym in Spanish) mandated the use of electronic invoices for all firm-to-firm sales. Our dataset encompasses all invoices dated from January 2019 to December 2023, and was obtained through the Central Bank of Chile.[5]

The invoices comprise two main sections: the "Heading" and the "Detail". The Heading section includes essential information regarding the invoice, such as the date, tax identifiers for both the seller and the buyer, the municipalities where they are located, and the payment terms (cash or credit). The Detail section contains specific information about the products sold. Each product is represented by an individual entry within this section. Entries include a description which varies in the level of detail (e.g. "Maule Graphics GC1 315GR 77X110 2500 PL" vs. "Metal plaque") and a product code (e.g. "ED003C3V59025A")[6] which identify the items being traded. Additionally, entries include price, quantity, and any applicable discounts or surcharges for each product.[7] For 90% of the sales, we have information on the units in which quantities are denominated. Finally, the Central Bank of Chile assigns each firm to a 3-digit ISIC Rev. 4 sector, and each product to broad categories according to the Central Product Classification (CPC).

[5]This paper 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 individuals or firms. Officials of the CBC processed the disaggregated data. All the analysis was implemented by the authors and did not involve nor compromise SII. The information contained in the databases of the SII 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.

[6]This product code corresponds to fully coated folding white boxboards used for the production of medical and healthcare packaging.

[7]For an invoice to be valid, the Chilean Internal Tax Service only requires the seller's and buyer's tax ID's, the text description of goods sold, the total amount paid, and a date. More information can be found (in Spanish) at https://www.sii.cl/factura_electronica/formato_dte.pdf.

In what follows, we use the terms "firm" to refer to a unique tax ID, and "establishment" to refer to a tax ID-municipality duplet. We treat all establishments of the same firm as a single seller or buyer. Our set of sellers includes all manufacturing firms, while our set of buyers includes all non-government operated firms in Chile.[8] Because we focus on firm-to-firm sales, we treat each retailer and wholesaler as a single buyer of the manufacturing products. Hence, our measures of price dispersion do not capture any variation in prices across the buyers of the manufactured products when resold by retailers and wholesalers.

We identify unique products using product descriptions and codes. These codes can take different forms, such as EAN (European Article Number) codes or internal codes used by the seller, and are only available to us starting in June 2021.[9] We focus on non-oil manufacturing products that are assigned codes consisting of at least four characters.[10] In our baseline analysis, we define a unique product as a product code-selling establishment combination. Thus, identical product codes issued by the same seller from different establishments are considered distinct products.[11] For each seller, we discard codes associated with multiple descriptions, as well as descriptions that are associated with multiple codes. We use both the descriptions and the codes to minimize concerns about descriptions being vague or internal product codes being coarse. We conduct robustness exercises using the smaller sample of products for which EAN product codes are available.

We take the following additional steps to clean the raw data. First, we address potential typos in the invoices by removing transactions that exceed 100 billion pesos (approximately 125 million USD) or that display non-positive prices or quantities. Second, we discard invoices in which the tax ID of the seller matches that of the buyer. Third, we eliminate invoices that were later cancelled or deemed invalid. Fourth, we merge the invoices with the firm's wage and tax statements and discard firms that do not declare any paid workers. Finally, to be conservative in our estimates of price dispersion, we eliminate transactions where the product's recorded price deviates from the average price recorded in other transactions from the same month by a ratio that is above exp(1)

[8] In our baseline sample of uniquely identified products, described below, 33% of firm-to-firm manufacturing sales are to other manufacturing firms, 51% are to firms Retail and Wholesale, and 16% are to firms in other sectors.

[9] The Central Bank of Chile has a data agreement with Chilean Internal Tax Service, which incorporates several tax forms and documents. Not all the data in the Chilean Internal Tax Service is part of this agreement.

[10] The median product in our sample consists of 8 characters.

[11] We observe the units in which quantities are denominated for 90% of the sales. When units are observed, all transactions of a given product have the same units. Our baseline sample includes products for which we do not observe units across all transactions, but our measures of price dispersion are fairly insensitive to dropping these products.

or below exp(-1). Appendix B describes these datasets and cleaning steps in detail.

2.2 Summary statistics

Our goal is to analyze price dispersion across buyers who purchase the same product within a narrow period of time. Our baseline analysis focuses on transactions occurring during June 2023.[12] Table 1 reports summary statistics for the sample of invoices issued by manufacturing firms in this month. These statistics do not vary much across months between June 2021 and December 2023.

The first column provides an overview of the entire manufacturing sample, consisting of 4.6 million invoices that amount to 5.0 billion USD in sales. The second column focuses on our baseline sample of 379,418 products that we can uniquely identify. This sample covers 2.1 million invoices and 10.5 million transactions, representing almost half of total manufacturing sales (2.4 USD billion). Even though the fraction of sellers using product codes is relatively small (4,345 out of 45,256), they account for almost half the sales and the invoices.

Products purchased by more than one buyer account for 55% of sales of our baseline sample. The average product in our sample had 9.8 buyers, with each buyer making repeat purchases of the same product 2.8 times on average during June 2023. When focusing on products with multiple buyers, the average number of buyers increases to 24.8, though the median of 4 buyers suggests a right-skewed distribution.

The third column of Table 1 reports summary statistics for the sample of products with EAN codes. These products constitute a small portion of the total sales (0.4 billion USD) and are used by only 220 sellers. Products with EAN codes are more frequently sold to multiple buyers and have a higher average number of buyers.

Table 2 shows industrial composition and sectoral inflation rates in our baseline sample and in the official manufacturing PPI data. The largest sector in our sample is 'Food', followed by 'Chemicals'. Importantly, sectorial shares in our data are very similar to those in the PPI: the correlation is 0.97. The level and cross-sectoral variation in inflation rates are also very correlated in our data and in the PPI data.

[12] We focus on this month as it was preceded by a year of relatively low PPI inflation in Chile. Specifically, manufacturing PPI inflation was -3.1% between June 2022 and June 2023.

Table 1: Summary statistics

	All	Baseline	EAN codes
Sales (USD bill.)	5.0	2.4	0.4
Invoices (mill.)	4.6	2.1	0.5
Transactions (mill.)	22.3	10.5	2.9
Sellers	45,256	4,345	220
Buyers	370,585	147,077	51,332
Products		379,418	43,177
Buyer-seller pairs		389,525	59,997
Buyer-product pairs		3,732,252	607,763
Buyers per-product		9.8	14.1
Transactions per buyer-product		2.8	4.8
Products with more than one buyer:			
Share of sales		0.55	0.67
Buyers per-product		24.8	26.5
Buyers for median product		4.0	4.0
Transactions per buyer-product		6.2	11.4
Transactions for median buyer-product pair		2.0	3.0

Notes: The table reports summary statistics for different sample of invoices. 'All' refers to all non-oil manufacturing invoices in June 2021. 'Baseline' refers to our baseline sample, which identifies unique products and also cleans the data as described in the text. 'EAN codes' refers to the sample of invoices where products can be identified using EAN codes.

3 Empirical findings

In this section we report our central empirical findings. We first present our baseline measures of price dispersion and analyze how it varies across time, sectors, products, and sellers. We then show that observed price gaps are not driven by transportation costs, mode of payment, or by staggered price changes across buyers. Next, we investigate how price gaps vary by characteristics of the buyer and the buyer-seller relation. Finally, we evaluate the importance of quantity discounts and assess whether our measure of price gaps is consistent with price/quantity menus.

Table 2: Comparison with official PPI data

	Share of sales		Inflation 06/21-06/22	
	Sample	PPI	Sample	PPI
Food	50.3	34.3	6.8	6.4
Beverages	6.9	10.3	11.3	7.2
Textiles	0.6	0.0	5.9	NA
Wood products	4.0	5.0	-2.0	-7.6
Paper	8.4	8.7	1.2	-20.7
Printing	0.1	1.7	12.4	9.7
Chemicals	10.3	11.2	-14.3	-15.7
Pharmaceutical	3.2	3.1	7.3	7.0
Plastic, rubber	3.7	4.3	-5.1	-6.9
Non-metallic minerals	4.9	4.8	10.7	9.3
Metals	3.6	2.6	-21.8	-9.5
Metal products	1.6	7.9	2.2	-0.3
Electric equipment	0.6	0.6	2.5	7.6
Machinery and equip	0.5	3.6	13.4	6.7
Furniture	1.0	1.9	2.1	5.1
Other	0.2	0.0	15.3	NA
Correlation	0.97		0.73	

Notes: The first panel reports the share of each manufacturing 2-digit ISIC sector in total 'non-oil' manufacturing sales. The first column reports the shares for the baseline sample, and the second column reports the shares used by the Chilean Institute of National Statistics for the computation of the Manufacturing Producer Price Index (PPI). The second panel reports the inflation between June 2022 and June 2023 for each 2-digit ISIC sector. The third column reports inflation in our sample, computed as a weighted average of log-price changes with weights based on sales shares in the invoices. The fourth column reports the official PPI inflation in each sector.

3.1 Price dispersion across buyers

Let $p_{ib\tau}$ denote the price (after discounts and surcharges) of product i purchased by buyer b in transaction τ.[13] Denote the price gap (in logs) for transaction τ by

$$\mu_{ib\tau} \equiv \log\left(p_{ib\tau}/p_i\right), \tag{1}$$

where p_i denotes some average of prices paid by buyers of product i. Our results in this section do not depend on the value of p_i.

We start by showing that for the typical product, most of the dispersion in price gaps within a month is across buyers rather than across transactions for the same buyer. To do so, we regress $\mu_{ib\tau}$ on a full-set of product-buyer fixed effects for the sample of products

[13] In our data, 20% of total sales receive a discount, 2% of sales have a surcharge, and 4% of our sales receive both a discount and a surcharge. We focus on final prices after discounts and surcharges as they correspond to the effective transaction prices.

purchased by multiple buyers during June 2023. Table 3 reports the partial R-squared of these regressions relative to a reduced model that only includes product-level fixed effects (no buyer fixed effects).[14] The partial R-squared ranges between 0.88 and 0.93 depending on the weighting schemes, showing that buyer fixed effects absorb a large portion of the variation in prices across transactions within the month.

Given that there is little variation in prices across transactions for a given product-buyer, we aggregate prices by product-buyer across transactions in each month "t" as $p_{ibt} \equiv [\sum_{\tau \in t} p_{ib\tau} q_{ib\tau}] \,/\, [\sum_{\tau \in t} q_{ib\tau}]$, and define the price gap for the ib pair as $\mu_{ibt} \equiv \log(p_{ibt}/p_{it})$, where p_{it} is an average of p_{ibt}.[15] Unless noted otherwise, we focus on $t =$ June 2023 and omit subscript "t" to streamline notation. We denote the vector of observed price gaps for different buyers of product i by $\boldsymbol{\mu_i}$.

Table 3: Price dispersion within and across buyers

	(1)	(2)	(3)
Partial R2	0.88	0.89	0.93
Observations	9,234,194	9,234,194	9,234,194
Weights	None	Sales	Quantity

Notes: The Table reports the partial R-squared of the model $\mu_{ib\tau} = FE_{ib} + \epsilon_{ib\tau}$ relative to the reduced model $\mu_{ib\tau} = FE_i + \epsilon_{ib\tau}$, where FE_i and FE_{ib} are a full set of product-level, and product-buyer level fixed effects. The regressions are estimated by OLS on the sample of products that were purchased by more than on buyer in June 2023. In Column 1, each transaction receives an equal weight. In Column 2 transactions are weighted by their sales. In Column 3, each product receives an equal weight, and transactions of the same product are weighted according to the transaction's share in the product's quantity.

We calculate two measures of price dispersion for multi-buyer products. The first is the range of price gaps across buyers (in log differences), $r_i \equiv \max(\boldsymbol{\mu_i}) - \min(\boldsymbol{\mu_i})$. The second is the weighted standard deviation of price gaps across buyers, $\sigma_i \equiv \text{stdev}(\boldsymbol{\mu_i})$.[16]

Table 4 presents the sales-weighted distributions of these two statistics across products. The first column reports the distribution of r_i, showing that buyers pay very different prices during the same month. The median (average) range is 0.32 (0.38), and the percentile 75 is 0.55. The second column displays the distribution of σ_i. For the median (average) product, the standard deviation is 0.08 (0.06), and the percentile 75 is 0.11.

For a sample of products (representing 92% of sales) we can link the CPC categories

[14] The partial R-squared equals $\left[1 - SSR^F/SSR^R\right]$, where SSR^F and SSR^R denote the sum of squared residuals of the full and the reduced model, respectively. This statistic gives the proportion of the variation explained by the full model that cannot be explained by the explanatory variables in the reduced model. Throughout this section, the reduced model used for the computations of the partial R-squared is a simple regression of price gaps and product fixed effects.

[15] Thus, $\mu_{ibt} = [\sum_{\tau \in t} \mu_{ib\tau} q_{ib\tau}] \,/\, [\sum_{\tau \in t} q_{ib\tau}]$, where $\mu_{ib\tau} = \log(p_{ib\tau}/p_i)$.

[16] The weighted standard deviation is computed as $\text{stdev}(\boldsymbol{\mu_i}) = \sqrt{\sum_b \mu_{ib}^2 \frac{q_{ib}}{\sum_b q_{ib}} - \left[\sum_b \mu_{ib} \frac{q_{ib}}{\sum_b q_{ib}}\right]^2}$.

assigned by the Central Bank of Chile to the product categories in the 8-digit Harmonized System (HS8) used by Chilean customs. The level of dissagregation of HS8 categories is similar to that used in previous work on price dispersion across buyers based on international trade unit value data.[17] Since CPC categories are typically coarser than HS8 categories, most products in our sample are concorded to multiple HS8 categories. However, we can concord a subset of these products (representing 15% of the sales), with a unique HS8 category. For this sample, the median σ_i is 0.54 if we define products as an HS8-selling establishment combination, while the median σ_i is 0.05 if we use our baseline product definition. This underscores the importance of using detailed product codes from invoices to measure price dispersion across buyers of the same products.

Table 4: Within-product standard deviation of price gaps: Robustness

	Baseline		Unweighted	Transactions	One year	EAN
	Range (r_i)	Stdev (σ_i)	σ_i	σ_i	σ_i	σ_i
P1	0.00	0.00	0.00	0.00	0.00	0.00
P5	0.00	0.00	0.00	0.00	0.00	0.01
P10	0.03	0.01	0.01	0.01	0.01	0.02
P25	0.15	0.03	0.04	0.03	0.03	0.05
P50	0.32	0.06	0.07	0.07	0.07	0.08
P75	0.55	0.11	0.12	0.12	0.11	0.13
P90	0.79	0.16	0.18	0.17	0.17	0.19
P95	0.97	0.21	0.24	0.22	0.23	0.23
P99	1.28	0.34	0.41	0.35	0.52	0.37
Mean	0.38	0.08	0.09	0.08	0.09	0.10

Notes: r_i is the difference between the highest and lowest price gap for product i. σ_i is the sales-weighted standard deviation of price gaps across buyers of product i. The columns under baseline reports the distributions of r_i and σ_i, where the unit of observation is a product i, and products are weighted by their sales and the prices are aggregated at the buyer level and include all June 2023 transactions; $p_{ib} = [\sum_\tau p_{ib\tau} q_{ib\tau} / \sum_\tau q_{ib\tau}]$ for $\tau \in$ June 2023. Column 'Unweighted' uses unweighted standard deviations. Column 'Transactions' uses transaction level price gaps $\mu_{ib\tau}$ instead of aggregating at the buyer level. Column 'One year' aggregates prices at the buyer level using all transactions between January and December 2023. Column 'EAN' uses the sample of products that can be identified by an EAN code. The distributions only include products with multiple buyers. P1-P99 represent percentiles of the distribution. The distributions only include products with multiple buyers. P1-P99 represent percentiles of the distribution.

[17] For example, Alviarez et al. (2023) uses HS10 categories in the US, Fontaine et al. (2020) use CN8 Categories in France, Ignatenko (2023) uses HS8 categories in Paraguay.

3.1.1 Robustness

In Table 4 we recalculate dispersion of price gaps for different choices of samples and specifications. Column 'Unweighted' reports the unweighted distribution instead of the sales-weighted one. Column 'Transactions' is based on price gaps by transaction, rather than first aggregating transactions by product-buyer. Notably, these distributions closely resemble the one in our baseline specification.

Column 'One year' uses product-buyer level prices calculated by aggregating prices across all transactions in the period from January to December 2023. The share of sales of products purchased by multiple buyers is larger (0.61 vs. 0.55) if we extend the period to a year, as different buyers purchase the same product in different months. However, for products with multiple buyers, the dispersion of price gaps over the year is similar to that in June 2023.

Finally, Column 'EAN' reports results for the smaller sample of products identified by an EAN code. Recall from Table 1 that these products have a greater number of buyers. Price dispersion in this sample is slightly higher than in our baseline sample.

3.1.2 Price dispersion and inflation

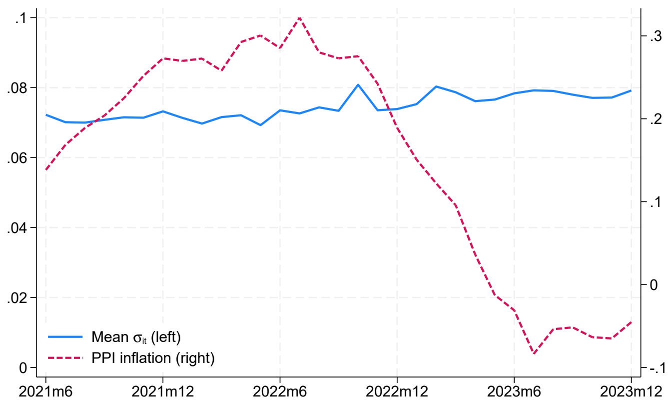
We now evaluate the extent to which price dispersion changes with the overall level of inflation, which displays large swings in our sample period. To do so, we compute the standard deviation of price gaps across buyers, σ_{it}, for every product i and month t between June 2021 and December 2023. Figure 1 plots σ_{it} for the average product in each month, and compares it to the year-to-year change in the PPI. The average standard deviation of price gaps across products changes very little over this period, increasing steadily from 0.07 in 2021 to 0.08 in 2023.[18] In contrast, annual PPI inflation rises from 13.8% in June 2021 to 28.5% in June 2022, and falls to -3.1% in June 2023. Appendix Figure A.1 extends this analysis to the sectoral level, showing that there is no systematic relation between inflation rates and average price dispersion by 2-digit ISIC sectors over time.

3.1.3 Which products have more dispersed prices?

We now study how the degree of price dispersion relates to product characteristics. We start by examining variation across manufacturing sectors and across the seller size distribution. Appendix Table A.1 reports the distributions of σ_i for each 2-digit ISIC sector, showing that price dispersion is pervasive across all sectors. The sectors with the most

[18]We obtain similar results if we use median rather than average σ_i.

Figure 1: Price dispersion and inflation



Notes: The blue line (left) plots the standard deviation of price gaps, σ_i, for the average product in each month between June 2021 and December 2023. The red line (right) plots the year-to-year change in the producer price index over the same periods.

dispersed prices across buyers are 'Paper' and 'Pharmaceutical' (average σ_i of 0.10 and 0.09), while the sectors with least dispersed prices are 'Beverages' and 'Wood products' (average σ_i of 0.05).

To examine the relation between price dispersion and size, we fit regressions of the form

$$\sigma_i = \beta' \mathbf{X}_i + FE_{municipality(s)} + FE_{sector(s)} + \epsilon_i. \tag{2}$$

The unit of observation is a product and the dependent variable is the standard deviation of price gaps across buyers of the product. The regressor $\mathbf{X}_i$ considers different measures of the size of the product, while $FE_{municipality(s)}$ and $FE_{sector(s)}$ are fixed effects for seller's municipality and the seller's 3-digit ISIC sector.[19] All regressions consider the sample of products that are purchased by more than one buyer.

The specification in Column 1 of Table 5 includes only the seller's municipality and sector, yielding an R-squared of 0.08. Column 2 shows a positive relation between seller size (measured using employment) and price dispersion across buyers. Column column 3 uses product sales as a measure of size, while Column 4 uses number of buyers. Finally, Column 5 introduces all these variables simultaneously. While the estimated elasticities are all statistically significantly positive, the R-squared in column 5 is near 0.13. Hence, these product characteristics do not account for a large fraction of the variation in disper-

[19]A municipality is the smallest administrative subdivision in Chile. There are 345 municipalities in Chile, 32 of which are in Santiago. Our data covers 51 3-digit ISIC manufacturing sectors.

sion across products in our sample.

Table 5: Price dispersion and product type

	(1)	(2)	(3)	(4)	(5)
$\log \text{emp}_{s(i)}$		0.0075***			0.0050***
		(0.0008)			(0.0008)
$\log \text{sales}_{s(i)}$			0.0054***		0.0020***
			(0.0004)		(0.0005)
$\log \#\text{buyers}_{s(i)}$				0.0147***	0.0110***
				(0.0014)	(0.0016)
Seller's sector FE	Y	Y	Y	Y	Y
Seller's municipality FE	Y	Y	Y	Y	Y
R^2	0.0804	0.1001	0.1060	0.1205	0.1328
Observations	140,027	140,027	140,027	140,027	140,027

Notes: The table reports the results of estimating equation (2) by OLS. $\text{emp}_{s(i)}$ refers to the number of employees working for the seller of product i. $\text{sales}_{s(i)}$ and $\#\text{buyers}_{s(i)}$ respectively denote the value of sales and the number of buyers of product i in June 2021. Seller's sector FE are fixed effects for 51 3-digit ISIC manufacturing sectors covered in our data, and Seller's municipality FE are fixed effects for the 345 municipalities in Chile. Standard errors in parentheses are clustered at the seller level. * significant at the 10% level, **: significant at the 5% level, *** significant at the 1% level.

3.2 Price gaps and differences in costs of supplying different buyers

Variation in distance between buyers and sellers and in the mode of payment by the buyer (cash or credit) can induce differences in the cost of supplying different buyers, resulting in price dispersion across buyers. In this section we show that these two observable characteristics play a small role in accounting for the observed price gaps in our data.

3.2.1 Shipping costs

We start by regressing price gaps on distance:

$$\mu_{ib} = \mathbb{I}\left(\log \text{dist}_{ib} = 0\right) + \beta \log \text{dist}_{ib} + FE_{municipality(b)} + FE_i + \epsilon_{ib}. \tag{3}$$

Here, dist_{ib} is the kilometer distance between the establishment that sells product i and the buyer,[20] $\mathbb{I}\left(\cdot\right)$ is an indicator for whether the selling establishment and buyer are in the

[20]We compute the distance between the selling establishment and each buyer as kilometers between the capital cities of their respective municipalities. For buyer-seller pairs from the same municipality, we assign a distance of 1. The distance between the two closest municipalities is 1.3 kilometers. For buyers that

same municipality, $FE_{municipality(b)}$ is a fixed effect for the municipality of the buyer, and FE_i is a product fixed effect.

Column 1 in Table 6 shows that the coefficient on the indicator variable is negative: price gaps are on average 0.0062 lower when the seller and the buyer are in the same municipality. The coefficient on distance in Column 2 is positive, indicating that the price is increasing in the distance between seller and buyer.

While the coefficients on distance are statistically significant, their economic impact is minimal. We illustrate this in a several ways. For the average product, the log difference in distance between the nearest and the farthest buyer is 1.3 (3.7 km),, which, based on the estimates in Column 2, corresponds to only a 0.003 log difference in prices. The largest log distance between two establishments in our sample is 8.4 (4,624 km), which implies a 0.017 higher price relative to a buyer that is just 1 km away from the seller. This price gap is small relative to the cross-buyer range in prices of 0.32 for the median product shown in Table 4. Moreover, if we exclude from prices any additional surcharges listed in the invoices, the correlation between prices and distance is not statistically significant. This suggests that shipping fees are reflected in the surcharges reported in the invoices.[21]

The partial R-squared for distance in regression (3), displayed in columns 1 and 2 of Table 6 is roughly 0.034. In a more flexible specification that includes quadratic terms and allows for 3-digit-ISIC sector-specific distance effects the partial R–squared rises to 0.041, indicating that distance does not account for much of the observed dispersion in price gaps across buyers. Finally, price dispersion using prices residualized by distance based on the flexible specification, reported in Column 2 in Table 7, is very similar to that in our baseline. In sum, shipping costs are not a significant driver of price dispersion across buyers in our data.[22]

3.2.2 Payment terms

To examine the extent to which price gaps reflect differences in payment terms, we leverage information on whether the invoice is paid in cash or credit. This information is

purchase the same product from multiple establishments (i.e. unique tax ID's appearing on the invoices purchasing a given product from different municipalities), we use the distance to the average establishment (weighted by purchases of the product). Such product-buyer pairs only comprise 10% of purchases in our sample.

[21] As noted in Section 3.1, our baseline results use prices inclusive of surcharges as these are the effective prices that buyers face. Our other empirical results in Section 3.1 and quantitative results in Section 4 are unchanged if we focus on prices exclusive of surcharges.

[22] The standard deviation of prices is also very similar in a sample of invoices in which both the seller and the buyer are located within the city of Santiago. These results are available upon request.

Table 6: Price gaps and differences in costs of supplying different buyers

	(1)	(2)	(3)
$\mathbb{I}(\log \text{dist}_{ib} = 0)$	-0.0062***	0.0011	
	(0.0013)	(0.0016)	
$\log dist_{ib}$		0.0020***	
		(0.0003)	
$\mathbb{I}(\text{credit}_{ib})$			-0.0194***
			(0.0009)
Product FE	Y	Y	Y
Buyer mun. FE	Y	Y	N
Partial R^2	0.0339	0.0341	0.0043
Observations	3,458,487	3,458,487	2,169,176

Notes: Columns 1 and 2 report the results of estimating equation (3) by OLS, and standard errors in parentheses are clustered at the buyer-seller level. Column 3 reports the results of estimating equation (4) by OLS, and standard errors in parentheses are clustered at the buyer-seller level. All columns include product-level fixed effects: * significant at the 10% level, **: significant at the 5% level, *** significant at the 1% level.

available for 62% of transactions, which account for 71% of the sales. Among the transactions for which data on payment terms are available, 69% percent (93% of the sales) are paid in credit.

We consider a regression of price gaps on mode of payment:[23]

$$\mu_{ib} = \mathbb{I}(\text{credit}_{ib}) + FE_i + \epsilon_{ib}, \tag{4}$$

where $\mathbb{I}(\text{credit}_{ib})$ is a dummy indicating that buyer b of product i paid in credit.

Column 3 of Table 6 shows that the coefficient on the indicator variable is negative: price gaps are on average 0.0194 lower when the buyer pays in credit. While this may seem counterintuitive, it is partly accounted by the fact that credit is used for bigger transactions and bigger transactions display lower price gaps (we examine this relation in Section 3.4). More importantly, the magnitude of the coefficient is small relative to the observed variation in prices, indicating that the payment terms do not account for much of the observed price gaps. In a more flexible specification that allows for sector-specific effects of payment method, the partial R-squared is only 0.011.[24] Finally, Column 3 in Table 7 shows that the dispersion in prices across buyers barely changes when prices are residualized by payment method using this flexible specification.

[23]We focus on the product-buyer pairs that always exhibit only one payment-term during the month. Buyers that use both cash and credit when purchasing the same product account for only 0.7% of sales.

[24]The full model in this specification is $\mu_{ib} = \sum_s [\mathbb{I}_{s,ib}(\text{credit})] + FE_i + \epsilon_{ib}$, where s is the sector of product i for the sectors used in Table 2.

Table 7: Residualized price dispersion

	Baseline	Residualized distance	Residualized payment
P1	0.00	0.00	0.00
P5	0.00	0.01	0.00
P10	0.01	0.01	0.01
P25	0.03	0.03	0.02
P50	0.06	0.06	0.06
P75	0.11	0.11	0.11
P90	0.16	0.16	0.17
P95	0.21	0.20	0.21
P99	0.34	0.34	0.34
Mean	0.08	0.08	0.08

Notes: The table reports the distribution of the standard deviation of prices across buyers of the same product for different samples. The unit of observation in these distributions is a product i. Column 'Baseline' uses our baseline sample as in Table 4. Column 'Residualized distance' reports the dispersion in the residuals from a regression like (3) that includes quadratic terms and allows for sector-specific distance effects. Column 'Residualized payment' reports the dispersion in the residuals from the regression in residuals from the regression like (4) that includes sector-specific dummies for whether the buyer pays in credit. The distributions only include products with multiple buyers. P1-P99 represent percentiles of the distribution.

3.3 Correlation of price gaps over time

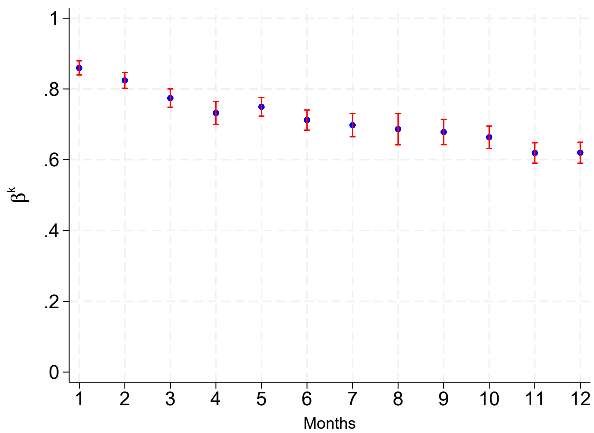
We now evaluate whether price gaps are persistent over time. To do so, we regress the price gap for buyer b, product i and month $t =$ June 2023 on the price gap for the same buyer and product in month $t - k$, controlling for product fixed effects:

$$\mu_{ibt} = \beta_k \mu_{ibt-k} + FE_i + \epsilon_{ib}. \tag{5}$$

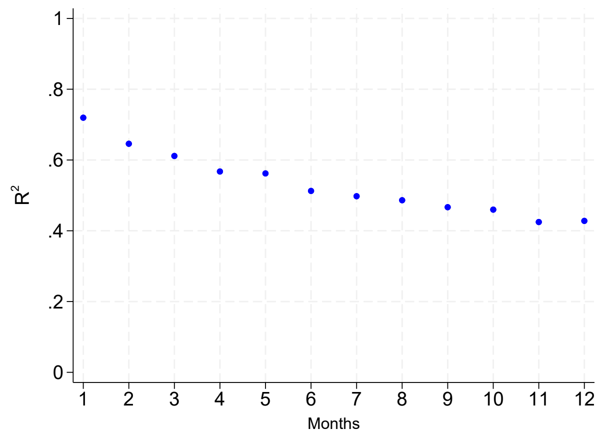
We estimate (5) for lags ranging from $k = 1$ to 12 months. The coefficient β_k measures the sensitivity of the price gap for a buyer-product in period t to the price gap k months before. The left panel of Figure 2 shows that price gaps are very persistent: β_k is above 0.8 for $k = 1$ and 0.6 for $k = 12$. The right panel shows that the partial R-squared is 0.72 for $k = 1$ and 0.43 for $k = 12$. Hence, past levels of price gaps account for most of the variation in current price gaps.

We obtain very similar results in a specification where μ_{ibt} and μ_{ibt-k} are calculated over the subset of product-buyer pairs that exhibit non-zero price changes between t and $t - k$, reported in Figure A.2 of the Appendix. This suggests that the persistence in price gaps is not driven mechanically by nominal price stickiness.

Figure 2: Persistence in price gaps



(a) Estimated coefficients β_k



(b) Partial R-squared

Notes: The left panel plots the coefficients β^k from equation (5) estimated by OLS, for $k = 1, ..., 12$. The regression includes product fixed effects. The red lines represent 95% confidence intervals. The right panel plots the corresponding partial R-squared of the regressions. The corresponding regressions results are reported in Appendix Table A.2.

3.4 What drives variation in price gaps?

This section evaluates how price gaps correlate with observable characteristics of buyers and of the buyer-seller pairs. We begin by showing that price gaps are highly correlated across products purchased by the same buyer, and evaluate how price gaps relate to buyer's size. Next we show that half of the variance in price gaps can be linked to the identity of buyer-seller pairs, and relate price gaps to observable characteristics of the buyer-seller relation. Finally, we show that buyers that purchase larger quantities of the same product pay lower prices, though quantity discounts do not account for much of the variation in price gaps.

3.4.1 Price gaps and buyer size

Do buyers face similar price gaps for different products they purchase? We start by estimating a simple regression of μ_{ib} on product fixed effects and buyer fixed effects for the sample of products with multiple buyers. Column 1 of Table 8 shows that the partial R-squared of this regression is 0.22, indicating that over one fifth of the observed dispersion in price gaps can be accounted for by the identity of the buyers.

To evaluate the relation between price gaps and observable buyer characteristics, we fit regressions of the form:

$$\mu_{ib} = \beta \log size_b + FE_{municipality(b)} + FE_{sector(b)} + FE_i + \epsilon_{ib}, \tag{6}$$

where $FE_{municipality(b)}$ and $FE_{sector(b)}$ and are fixed effects for the buyer's municipality and 3-digit ISIC sector.[25]

Column 2 of Table 8 includes fixed effects for the buyer's sector and municipality, in addition to product fixed effects, but does not include buyer size. The partial R-squared for the sector and municipality fixed effects is 0.041, which is roughly one fifth of the partial R-squared of the buyer fixed effects. Columns 3 and 4 adds two measures of buyer size: total sales and purchases of intermediate inputs. Larger buyers pay lower prices.[26] The partial R-squared, however, is only marginally larger than in Column 2. In a more flexible specification that further includes quadratic terms and allows the coefficients to vary across each of the 51 manufacturing ISIC industries, the partial R-squared only increases to 0.065.[27] This indicates that, whereas the buyer identity explains a sizable share of price gaps, buyer size, sector, and municipality do not explain much of these price gaps.

3.4.2 Price gaps and buyer-seller relations

Are price gaps correlated across products within buyer-seller pairs, and which buyer-sellers characteristics account for these gaps? Column 5 of Table 8 fits a regression of μ_{ib} on product fixed effects and buyer-seller fixed effects. In order to be included in this regression, the buyer-seller pair must transact in more than one product. The partial R-squared of this regression is 0.42, indicating that buyer-seller dummies account for a sizable share of the residual variation in price gaps.

To evaluate the relation between price gaps and observable characteristics of buyer-seller pairs, we consider regressions of the form

$$\mu_{ib} = \beta' \mathbf{X}_{s(i)b} + FE_{sector(b)} + FE_{municipality(b)} + FE_i + \epsilon_{ib}, \tag{7}$$

where $\mathbf{X}_{s(i)b}$ contains observable characteristics of the relation between the buyer b and the seller of product, including buyer's municipality and 3-digit ISIC sector fixed effects.

We first evaluate how price gaps vary with the length of the relation between the seller

[25] There are buyers in 158 different 3-digit ISIC sectors.

[26] We obtain similar results if we measure size by the buyer's employment.

[27] If we allow the coefficients to vary across 134 CPC manufacturing product categories, the partial R-squared is 0.066.

Table 8: Price gaps, buyer size, and characteristics of the buyer-seller relation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log sales_b			-0.0018***				
			(0.0005)				
log purch_b				-0.0030***			
				(0.0006)			
$0 < \text{relation}_{s(i)b} < 1$ yr						-0.0040***	
						(0.0010)	
$\text{relation}_{s(i)b} \geq 1$ yr						-0.0086***	
						(0.0009)	
log $sales_{s(i)b}$							-0.0058***
							(0.0003)
Product FE	Y	Y	Y	Y	Y	Y	Y
Buyer FE	Y	N	N	N	N	N	N
Buyer X Seller FE	N	N	N	N	Y	N	N
Buyer's mun. FE	N	Y	Y	Y	N	Y	Y
Buyer's sector. FE	N	Y	Y	Y	N	Y	Y
Partial R2	0.224	0.0414	0.0423	0.0435	0.424	0.0430	0.0492
Observations	3,444,143	3,444,143	3,444,143	3,444,143	3,382,932	3,382,932	3,382,932

Notes: Column 1 reports the results of a regression of log price on product and buyer fixed effects. Columns 2-4 report the results of estimating equation (6) by OLS. sales_b and purch_b respectively denote the sales and intermediate input purchases of buyer b. Column 5 reports the results of a regression of log price on product and buyer X seller fixed effects. Columns 6-7 report the results of estimating equation (7) by OLS. "$0 < \text{relation}_{s(i)b} < 1$ yr" and "$\text{relation}_{s(i)b} \geq 1$ yr" respectively denote indicator variables for whether the first recorded transaction in the data for the corresponding buyer-seller pair occurred over June 2022-May 2023 (i.e. repeated relations as of June 2023, which started over the previous 12 months), and before June 2022. $sales_{s(i)b}$ are the total sales of the seller of product i to buyer b. Buyer's sector FE are fixed effects for 158 3-digit ISIC sectors covered in our data. Standard errors in parentheses are clustered at the buyer level * significant at the 10% level, **: significant at the 5% level, *** significant at the 1% level.

and the buyer. We define the length of the relation between buyer b and the seller of product i as the number of months from their first transaction since January 2019, and construct dummies indicating whether the relation is new (no previous transactions), started over the prior year, or is longer than a year. Column 6 in Table 8 shows a regression of price gaps on these dummies, where the omitted category is new relations. Price gaps are 0.004 (0.0086) log points lower for repeated relations that are less than a year old (repeated relations that are more than a year) compared to new relations. Next, column 7 displays the relation between price gaps and total sales of the seller of product i to buyer b. Price gaps are lower for buyer-seller pairs with higher sales.

The partial R-squared reported in Columns 6 and 7 is roughly 0.043, which is similar to the R-squared reported in Column 2 that only includes buyer's sector and municipality dummies. In a regression that simultaneously includes relation length, sales (linear and quadratic), and allows for sector-specific coefficients, the partial R-squared increases only to 0.05. Therefore, whereas observed buyer-seller characteristics are systematically correlated with price gaps, they only account for a small fraction of the differences price gaps across buyer-seller pairs.

3.4.3 Price gaps and quantity discounts

Are price gaps related to quantity purchased? We pool data for all months between June 2021 and December 2023 to lever variation across buyers and time and estimate regressions of the form:

$$\mu_{ibt} = \beta \log q_{ibt} + FE_{it} + \epsilon_{ibt}, \tag{8}$$

where q_{ibt} is the quantity of product i purchased by buyer b in month t, and FE_{it} is a product-month fixed effect.

Column 1 of Table 9 shows that buyers purchasing larger quantities of a product face lower price gaps, with an estimated elasticity of -0.015. For the median product, the observed range in log quantities purchased by different buyers is 1.95 log points, implying difference in log prices of only 0.03. We obtain very similar results if we limit our sample to June 2023.

Columns 2 and 3 include buyer-seller-month and buyer-product fixed effects to examine the relation between price-gaps and quantities across products within buyer-seller relations, and across time within buyer-products. Larger quantities are associated with lower prices along both dimensions, though the elasticity on quantity drops to -0.011 if we include buyer-seller-months fixed effects, and to -0.006 if we include buyer-product fixed effects.

The partial R-squared associated to quantities is very low (0.029 in column 1 of Table 9 and even lower in columns 2 and 3). In more flexible specifications that include quadratic terms and allows for sector-specific quantity coefficients, the partial R-squared associated to quantity is still under 0.04. That is, quantities do not explain much of the residual variation in price gaps.

3.5 Price-quantity menus

In Section 4 we will quantify the productivity losses arising from variation in prices across buyers. We will make the assumption that buyers can purchase any desired quantity of an input at the transacted price. An alternative assumption is that sellers engage in second-

Table 9: Price gaps and quantities

	(1)	(2)	(3)
$\log q_{ibt}$	-0.0149***	-0.0112***	-0.0062***
	(0.0004)	(0.0001)	(0.0001)
Product X time FE	Y	Y	Y
Buyer X seller X time FE	N	Y	N
Buyer X product FE	N	N	Y
Partial R-squared	0.0293	0.0128	0.0045
Observations	95,670,271	95,670,271	95,670,271

Notes: The Table reports the results of estimating equation (8) by OLS. All columns include product-time fixed effects. For the partial R-squared, the reduced model includes all the fixed effects indicated in the corresponding column. Standard errors in parentheses are clustered at the buyer-seller level. * significant at the 10% level, **: significant at the 5% level, *** significant at the 1% level.

or third-degree price discrimination, offering menus of quantities and prices. In this case, prices are not allocative other than by making certain menus more attractive to some buyers, and price gaps are not informative of misallocation.

The assumption of fixed menus implies a one-to-one mapping between quantities and prices. One testable implication of this assumption is that variation in quantities should entirely account for variation in prices. The low partial R-squared associated to quantity in the regressions discussed above already suggests that this is not the case. However, changes in prices and quantities may reflect other shocks that make producers change the menu they offer.

To test the hypothesis of non-linear prices set by sellers to individual buyers we examine whether, across transactions of a particular product-buyer pair, a given price is associated to a constant quantity. Specifically, we first compute for each product-buyer pair the standard deviation of log prices and log quantities across transactions that occurred over a certain time period. The first column in Table 10 considers transactions in June 2023 and reports the standard deviation of quantities for product-buyer pairs that have a standard deviation of prices that lower than 1%. If quantity-price menus are prevalent, we should observe a small dispersion of quantities across observations. Instead, we observe substantial dispersion in quantities among these buyers that consistently pay the same price across transactions.

The second, third, and fourth columns of Table 10 report log-changes in quantities for each product-buyer pair between May and June 2023, between March and June 2023, and between December 2022 and June 2023, considering only product-buyer pairs that experienced a price change smaller than 1% during each of these periods. Once again, we observe large changes in quantities associated with product-buyer pairs with small price movements. These results suggest that fixed price-quantity menus between buyers and sellers are not prevalent in our data. Of course, this evidence is not inconsistent with price discrimination by producers across sellers, which is the focus of the paper.

Table 10: Quantity-price menus

	$\text{stdev}_{ib}(\log q_{ib\tau})$	$\lvert \log q_{ibt} - \log q_{ibt-1} \rvert$	$\lvert \log q_{ibt} - \log q_{ibt-3} \rvert$	$\lvert \log q_{ibt} - \log q_{ibt-6} \rvert$
p1	0.00	0.00	0.00	0.00
p5	0.00	0.01	0.01	0.02
p10	0.01	0.04	0.04	0.06
p25	0.21	0.12	0.16	0.18
p50	0.44	0.30	0.39	0.41
p75	0.63	0.69	0.77	0.83
p90	0.85	1.25	1.41	1.57
p95	1.00	1.79	1.95	2.35
p99	1.54	3.22	3.30	4.67
Mean	0.45	0.52	0.60	0.69

Notes: The first column reports the distribution of the standard deviation of quantities across transaction for a given product-buyer pair, in which the standard deviation of prices is less than 0.01. The distribution excludes product-buyer pairs that record only one transaction. The second, third, and fourth column respectively report the distribution of the log-change in quantities for each product-buyer pair between May and June 2023, between March and June 2023, and between December 2022 and June 2023 for pairs that experienced a price change of less than 1% during these periods. The unit of observation in these distributions is a product-buyer pair *ib*. The distributions are sales weighted.

3.6 Taking stock

The results from the last two sections reveal a novel set of facts about how prices of intermediate inputs vary across buyers. First, there is substantial dispersion in the prices that different buyers pay for the same input within the same month. This dispersion in prices is very stable across months in our sample (featuring large swings in the average inflation rate), is pervasive across manufacturing sectors, and is more pronounced for products with higher sales, sold by larger sellers, and purchased by a larger number of buyers.

Second, observed price gaps are highly persistent over time (even conditional on a nominal price change) and strongly correlated across different products purchased by the same buyer. Prices are lower for larger buyers, for buyer-seller pairs with older relations and higher volume of sales, and are decreasing in the volume of purchased quantities. However, these observed characteristics of the buyer, seller, and transacted quantities account for a small fraction of the variation in price gaps. We also establish that shipping fees and form of payment play a small role in accounting for observed price gaps. The latter observation motivates us to interpret price gaps across buyers as differences in markups rather than as differences in the cost of supplying different buyers.

In what follows, we aim to quantify the productivity losses arising from price dispersion in the production network. We do not model the endogenous determination of markups in the observed allocations. Any model of endogenous markups would have to confront the observation that buyer-seller relationships account for a significant portion of the measured price gaps, but observed characteristics (such as buyer or seller size and transacted quantity) do not. Finally, we assume that the price gaps are allocative and that buyers can purchase any desired quantity of an input at the observed prices, which is consistent with the lack of evidence of quantity-price menus.

4 Model

In this section we describe the model that we use to quantify the aggregate productivity losses resulting from markup dispersion across buyers.

Preliminaries: We consider a closed economy with J sectors indexed by j, each populated by a discrete number of firms. Firms produce differentiated products, indexed by i, using a a bundle of intermediate inputs and services from a factor that is in fixed aggregate supply. Firms sell their products to other firms and to final consumers. Markups are exogenous and can differ across products and buyers.

Technologies of final goods: Final consumption C aggregates output of different sectors according to

$$C = \left[\sum_{j \in J} \left[\bar{\gamma}^j \right]^{\frac{1}{\eta}} \left[C^j \right]^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}} .$$

Sectoral output C^j aggregates output of products in sector j according to

$$C^j = \left[\sum_{i \in N_j} \bar{\gamma}_i^{\frac{1}{\rho}} c_i^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} ,$$

where c_i denotes quantity of product i in final output and N_j is the set of sector j products.

Technologies of intermediate products: The production function for producer i is:

$$y_i = z_i \left[[1-\bar{\alpha}_i]^{\frac{1}{\sigma}} l_i^{\frac{\sigma-1}{\sigma}} + \bar{\alpha}_i^{\frac{1}{\sigma}} m_i^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

Here l_i denotes the use of factor services by producer i, and m_i is a bundle of intermediate inputs:

$$m_i = \left[\sum_{j \in J} \left[\bar{\omega}_i^j\right]^{\frac{1}{\eta}} \left[m_i^j\right]^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}},$$

where the sectoral j bundle of intermediate inputs used by producer i is

$$m_i^j = \left[\sum_{i' \in N_j} \bar{\omega}_{i'i}^{\frac{1}{\rho}} x_{i'i}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

and $x_{i'i}$ denotes the quantity of product i' used in the production of i.

Market clearing: Goods market clearing implies:

$$y_i = c_i + \sum_{i'} x_{ii'} \qquad \forall i, \tag{9}$$

and factor market clearing implies

$$\sum_i l_i = \bar{L}. \tag{10}$$

Prices and wedges: Let $p_{ii'}$ and and p_{ic} respectively denote the price of product i when sold to i' and when sold to final consumers. These prices are given by:

$$p_{ii'} = \mu_{ii'} mc_i; \quad p_{ic} = \mu_{ic} mc_i, \tag{11}$$

where μ_{ii} and μ_{ic} are exogenous markups and mc_i is the marginal cost of production of product i. We let the factor price be the numeraire.

Equilibrium: Given technologies and wedges $\{\mu_{ii}, \mu_{ic}\}$, an equilibrium is a set of goods prices $\{p_{ii'}, p_{ic}\}$, intermediate input quantities $\{x_{ii'}\}$, factor quantities $\{l_i\}$, outputs $\{y_i\}$, and final output quantities $\{c_i\}$, such that (i) firms minimize costs and set prices to $p_{ii'} = \mu_{ii'} mc_i$ and $p_{ic} = \mu_{ic} mc_i$; (ii) final producers minimize costs; and (iii) all markets clear. Appendix C characterizes the equilibrium.

4.1 System in changes

We now provide a system of equations to evaluate how equilibrium prices and quantities change in response to exogenous changes in markups. Denote the ratio of a variable in the new relative to the initial equilibrium by $\hat{X} \equiv \frac{X'}{X}$. The changes in product level prices can be expressed as:

$$\hat{p}_{ii'} = \hat{\mu}_{ii'}\widehat{mc}_i; \quad \hat{p}_{ic} = \hat{\mu}_{ic}\widehat{mc}_i, \tag{12}$$

with

$$\widehat{mc}_i = \left[1 - \alpha_i + \alpha_i \hat{q}_i^{1-\sigma}\right]^{\frac{1}{1-\sigma}}. \tag{13}$$

Here, $\alpha_i \equiv \frac{q_i m_i}{l_i + q_i m_i}$ is the cost share of materials for product i in the initial equilibrium, and $\hat{q}_i$ is the change in the price of the input bundle used in the production of i:

$$\hat{q}_i = \left[\sum_j \omega_i^j \left[\hat{q}_i^j\right]^{1-\eta}\right]^{\frac{1}{1-\eta}}, \quad \text{where} \quad \hat{q}_i^j = \left[\frac{1}{\omega_i^j}\sum_{i' \in j} \omega_{i'i}\hat{p}_{i'i}^{1-\rho}\right]^{\frac{1}{1-\rho}}, \tag{14}$$

with $\omega_{i'i} \equiv \dfrac{p_{i'i}x_{i'i}}{q_i m_i}$ and $\omega_i^j \equiv \sum_{i' \in j} \omega_{i'i}$. Changes in final good prices are given by:

$$\hat{P} = \left[\sum_j s^{jc} \left[\hat{P}^j\right]^{1-\eta}\right]^{\frac{1}{1-\eta}}, \quad \text{where} \quad \hat{P}^j = \left[\frac{1}{s^{jc}}\sum_{i \in j} s_i^c \hat{p}_{ic}^{1-\rho}\right]^{\frac{1}{1-\rho}}, \tag{15}$$

with $s_i^c \equiv \frac{p_{ic}c_i}{\sum_i p_{ic}c_i}$ and $s^{jc} \equiv \sum_{i \in j} s_i^c$. Finally, goods market clearing implies:

$$\hat{y}_i = \frac{c_i}{y_i}\hat{p}_{ic}^{-\rho}\left[\hat{P}^j\right]^{\rho-\eta}\hat{P}^{\eta}\hat{C} + \left[1 - \frac{c_i}{y_i}\right]\sum_{i'} s_{ii'}\hat{p}_{ii'}^{-\rho}\left[\hat{q}_{i'}^j\right]^{\rho-\eta}\hat{q}_{i'}^{\eta-\sigma}\widehat{mc}_{i'}^{\sigma}\hat{y}_{i'}, \qquad \forall i \tag{16}$$

where $s_{ii'} \equiv x_{ii'} / \sum_{i'} x_{ii'}$. Factor market clearing implies

$$1 = \sum_i s_i^l \widehat{mc}_i^{\sigma}\hat{y}_i. \tag{17}$$

with $s_i^l \equiv l_i / L$.

The system of equations listed above can be used to solve for changes in aggregate consumption $\hat{C}$, which coincides with the change aggregate productivity given that factor quantities are fixed in our counterfactuals. To solve the system we need to assign values to the cost share of inputs, α_i, the share of each input i' in input expenditures, $\omega_{i'i}$, the share of the fixed factors used for the production of each product, s_i^l, the share of each product in final sales, s_i^c, and the fraction of the product that is sold to final consumers and to each other intermediate product i', $\frac{c_i}{y_i}$ and $s_{ii'}$. We also need to assign values to the elasticities of substitution between factors and materials σ, and the elasticities of substitution across products and sectors, ρ and η. We describe our calibration strategy in Section 5.

4.2 Analytic results under simplified network

To provide some intuition on how reducing price dispersion increases aggregate productivity, consider the following simplified network. Assume that there are two sectors labeled u and d (upstream and downstream). Goods in sector u are produced using factor services and are sold to the d sector: $\alpha_i = 0$ and $c_i/y_i = 0$ for $i \in N_u$. Conversely, goods in sector d are produced using intermediate inputs and sold to final consumers: $\alpha_i = 1$ and $c_i/y_i = 1$ for $d \in N_d$.

Let $\mathcal{M}_i \equiv \sum_{i'} \mu_{ii'} s_{ii'}$ be the quantity-weighted average markup of product i in the initial equilibrium. To eliminate markup dispersion while keeping the average markup unchanged for intermediate goods $i \in N_u$, we set $\hat{\mu}_{ii'} = \frac{\mathcal{M}_i}{\mu_{ii'}}$. For downstream goods $i \in N_d$, we leave markups unchanged, $\hat{\mu}_{ic} = 1$.

In Appendix C we show that the change in aggregate consumption (which equals the change in aggregate productivity) is

$$\hat{C} = \left[\sum_{i \in N_u} s_i^x \frac{\overline{\mathcal{M}_u}}{\mathcal{M}_i} \sum_{i' \in N_d} \hat{\mu}_{ii'}^{-\rho} s_{ii'} \right]^{-1} \left[\sum_{i \in N_u} s_i^x \sum_{i' \in N_d} \frac{\mathcal{M}_{i'}}{\overline{\mathcal{M}_d}} \hat{\mu}_{ii'}^{-\rho} s_{ii'} \right]^{\frac{\rho}{\rho-1}}, \tag{18}$$

where $s_i^x \equiv \frac{\sum_{i'} p_{ii'} x_{ii'}}{\sum_{i \in N_u} \sum_{i'} p_{ii'} x_{ii'}}$ is the share of input i in total input sales and $\overline{\mathcal{M}_k}$ is the average markup in sector k. If across downstream firms markups on final sales are uncorrelated with markups on their purchases ($\mathcal{M}_{i'}$ uncorrelated to $\hat{\mu}_{ii'}$ across $i' \in N_d$), and across upstream firms average markups are uncorrelated with markup dispersion over different buyers ($\mathcal{M}_i$ uncorrelated to $\sum_{i' \in N_d} \hat{\mu}_{ii'}^{-\rho} s_{ii'}$ across $i \in N_u$), equation (18) simplifies to

$$\hat{C} = \left[\sum_{i \in N_u} s_i^x \xi_i \right]^{\frac{1}{\rho-1}}, \qquad \text{where} \quad \xi_i \equiv \frac{\sum_{i'} s_{ii'} \mu_{ii'}^{\rho}}{\left[\sum_{i'} s_{ii'} \mu_{ii'}\right]^{\rho}} \geq 1 \quad \text{if } \rho > 1. \tag{19}$$

In this case, aggregate productivity rises if markups are initially dispersed across buyers and $\rho > 1$. The rise in C is increasing in markup dispersion and in the value of the elasticity ρ.

Without imposing these assumptions, equation (18) shows that $\hat{C}$ depends also on the covariance between changes in wedges $\hat{\mu}_{ii'}$ and product-level markups $\mathcal{M}_{i'}$. Aggregate productivity can decline if, for example, initial markup dispersion is lower for downstream products with high average markups, as this shifts production towards these high-markup firms. In our quantitative exercise, we target the shares that characterize the solution to the system in Section 5, which in turn fixes the covariance between $\mathcal{M}_{i'}$ and $\hat{\mu}_{ii'}$.[28]

5 Quantitative results

In this section we describe how we map the model to the Chilean electronic invoice data, and we quantify the aggregate productivity gains from equating markups across buyers.

5.1 Mapping the model to data

5.1.1 Sample of products

We consider the network of products which can be identified uniquely as described in Section 2. We additionally include buyers of these products from the manufacturing, retail, and wholesale sectors. As summarized in Appendix Table A.3, in June 2023 there are 240,763 products and 123,111 firms in this network. This results in more than 30 billion possible product-buyer pairs, of which 3,018,833 have positive sales. Our counterfactual exercises thus require solving a large non-linear system of equations. In Appendix E we describe our solution algorithm.

5.1.2 Measuring dispersion in markups across buyers

If marginal costs of production are common across buyers, differences in markups across buyers can be measured from observed differences in prices (i.e. our measure of price gaps). Equalizing markups across buyers while keeping the average unchanged, the

[28] Specifically, equation (18) can be written as $\hat{C} = \left[\sum_{i\in N_u} s_i^l \sum_{i'\in N_d} s_{ii'} \hat{\mu}_{ii'}^{-\rho}\right]^{-1} \left[\sum_{i\in N_d} s_i^c \sum_{i'\in N_u} \omega_{i'i} \hat{\mu}_{i'i}^{1-\rho}\right]^{\frac{\rho}{\rho-1}}$, and our calibration strategy targets s_i^l, s_i^c, $\omega_{i'i}$, and $s_{ii'}$ as discussed in Section 5.

change in product-level markups is:

$$\hat{\mu}_{ii'} = \frac{\sum_b \mu_{ib} s_{ib}}{\mu_{ii'}} = \frac{\sum_b \mu_{ib} mc_i s_{ib}}{\mu_{ii'} mc_i} = \frac{\sum_{ib} p_{ib} s_{ib}}{p_{ii'}}. \tag{20}$$

We use our price gap data presented in Section 2 to compute (20) for the set of products with unique codes. Note that prices p_{ib} are available for all buyers b observed in the invoices, regardless of whether the buyer uses product codes when issuing invoices. With this in mind, we divide products into two sets (which also correspond to our two sectors $j \in J$): those that have product codes and can be identified as unique, $i \in S$, and those that do not, $i \notin S$. In our counterfactual exercises, we change markups for products $i \in S$ with multiple buyers according to (20), and keep markups for single-buyer products $i \in S$ and for products $i \notin S$ constant, $\hat{\mu}_{ii'} = 1$. We also keep markups for all final sales constant, $\hat{\mu}_{ic} = 1$.

Products in $i \in S$ account for 30% of total sales. Firm-to-firm sales of these products account for 9% of the total sales, and those with multiple buyers account for 5.2% of total sales.

5.1.3 Parameterization

To evaluate how changes in markups affect aggregate productivity using the system in Section 4.1, we need to assign values to the cost share of inputs, α_i, the share of each input i' in input expenditures, $\omega_{i'i}$, the share of the fixed factors used for the production of each product, s_i^l, the share of each product in final sales, s_i^c, and the fraction of the product that is sold to final consumers and to each other intermediate product i', $\frac{c_i}{y_i}$ and $s_{ii'}$. We also need to assign values to the elasticities of substitution between factors and materials σ, and the elasticities of substitution across products and sectors, ρ and η.

Rather than making distributional assumptions to calibrate the shares, we map each product in the model to a product in the Chilean data. A challenge arises due to the fact that, while most firms in the data sell multiple products, data on input use and on sales to final consumers are only available at the firm level. Furthermore, the invoice data identify the buyer of each input but not how buyers apportion inputs across their multiple products.[29] To apportion firm-level inputs to products, we assume that each firm uses the same input bundle and sets the same average markup across its multiple products. We also assume the ratio of sales to final consumers to total sales is the same across all products of the same firm. Appendix D shows that under these assumptions,

[29] That is, the invoices have data on prices and quantities at the product-buyer level p_{ib} and x_{ib}, where the buyer b is identified by its tax ID.

we can apportion firms' inputs across products in proportion to intermediate sales.

To calibrate the shares listed above, we complement the invoice data with income declaration forms collected by the Chilean tax authority (Form 22 or IDF), which provide firm-level data on input expenditures, total costs, and value added. We treat purchases of inputs that are not in the model's network as purchases of factor services.[30] We denote by θ_f the share of input purchases of firm f that are included in the model's network by θ_f, by $f(i)$ the identity of the firm that produces product i, and by $\kappa_i \equiv [\sum_b p_{ib} x_{ib}] \,/\, \left[\sum_{i \in f(i)} \sum_b p_{ib} x_{ib}\right]$ the share of product i sales in total sales of firm $f(i)$.

We set α_i as the ratio of domestic input expenditures to total costs of firm $f(i)$ from the income declaration form, and multiply it by the share of input purchases included in the network, $\theta_{f(i)}$. We set $\omega_{i'i}$ as the ratio of purchases of product i' relative to all input purchases by firm $f(i)$ from the invoices. We set $s_{ii'}$ as the product of the quantity share of product i that goes to firm $f(i')$, $\frac{x_{if(i')}}{\sum_{f'} x_{if'}}$, and the share of product i' in $f(i')$ sales, $s_{ii'} = \frac{x_{if(i')}}{\sum_{f'} x_{if'}} \times \kappa_i$.

We obtain the remaining shares after first calibrating the following auxiliary variables. First, from the invoices we compute the share of each product i in total intermediate input sales, s_i^x, and the share of each product in total intermediate input purchases, s_i^m.[31] Second, use the income declaration forms to compute the ratio of value added to revenues for each product i, which we denote by v_i. We then obtain the remaining shares as $s_i^c = \left[s_i^m \frac{1}{1-v_i} - s_i^x\right] \frac{1-v}{v}$, $s_i^l = \left[\frac{1-\alpha_i}{\alpha_i} / \frac{1-\alpha}{\alpha}\right] s_i^m$, and $\frac{c_i}{y_i} = \left[1 + \frac{1-v}{v} \frac{s_i^x}{s_i^c}\right]^{-1}$. Here, v and α denote the economy-wide ratio of value added to revenues and the economy-wide ratio of input purchases to total cost. Using the previously targeted shares, we obtain these ratios as $v = \left[\sum_i s_i^m / v_i\right]^{-1} = 0.76$ and $\alpha = \left[\sum_i s_i^m / \alpha_i\right]^{-1} = 0.27$. Note that these choices pin down the ratio of revenues to costs , $\frac{\text{revenues}_\text{i}}{\text{cost}_\text{i}} = \frac{\alpha_i}{1-v_i}$, which corresponds to the average markup of product i. The ratio of total revenues to total costs in the economy is $\frac{\alpha}{1-\nu}$ =1.13. Table 11 summarizes the calibration of initial shares, and Appendix D provides additional details.

Finally, we set the the elasticity of substitution between factor services and the input bundle to $\sigma = 1$, and assume that the elasticity of substitution between sectors equals the elasticity of substitution between intermediate goods, $\rho = \eta$. We report our findings for low and high values of ρ: $\rho = 3$ or $\rho = 6$.[32] We provide robustness of our quantitative

[30] As noted above, we focus on the network of products for which we observe product identifiers, and additionally include all the buyers of these products in both the Manufacturing and Retail and Wholesale sectors. This implies that, for example, purchases of agricultural inputs are not considered in the network and are lumped with into factor services.

[31] Specifically, we set $s_i^x \equiv [\sum_{i'} p_{ii'} x_{ii'}] \,/\, [\sum_i \sum_{i'} p_{ii'} x_{ii'}]$, $s_i^m = \left[\sum_{i'} p_{if(i)} x_{i'f(i)}\right] / \left[\sum_{f'} \sum_{i'} p_{i'f'} x_{i'f'}\right] \times k_i$.

[32] It's is well known that the cost of markup dispersion increases with the value of the elasticity of substitution. Hsieh and Klenow (2009) set this elasticity to 3, Baqaee and Farhi (2020) consider values of 4 and 8,

results to these assumptions.

Table 11: Calibration of initial shares

Targeted shares

α_i	Ratio of purchases of input to total cost of firm $f(i)$, IDF, EI.
$\omega_{i'i}$	Share of intermediate input i' in input purchases of firm $f(i)$, EI.
s_i^x	Share of input i in total intermediate input sales, EI.
κ_i	Share of product i in firm's $f(i)$ sales, EI.
$s_{ii'}$	Quantity share of intermediate of i going to $f(i')$ multiplied by κ_i, EI.
s_i^m	Share of $f(i)$ in total intermediate input purchases multiplied by κ_i, EI.
v_i	Ratio value added to revenues of firm $f(i)$. IDF, EI.

Non targeted shares

s_i^c	Share of sales product i in total final sales: $s_i^c = \left[s_i^m \frac{1}{1-v_i} - s_i^x\right] \frac{1-v}{v}$.
s_i^l	Share of the fixed factor that is used for the production of i: $s_i^l = \left[\frac{1-\alpha_i}{\alpha_i} / \frac{1-\alpha}{\alpha}\right] s_i^m$.
$\frac{c_i}{y_i}$	Fraction of good i that is sold to final consumers: $\frac{c_i}{y_i} = \left[1 + \frac{1-v}{v} \frac{s_i^x}{s_i^c}\right]^{-1}$.
α	Economy wide share of inputs in production cost: $\alpha \equiv \left[\sum_i s_i^m / \alpha_i\right]^{-1} = 0.27$
v	Economy wide ratio of value added to revenues. $v \equiv \left[\sum_i s_i^m / v_i\right]^{-1} = 0.76$

Notes: This table describes how we assign values to the model's initial shares under our baseline calibration. $f(i)$ refers to the firm (tax id) that sells product i. 'EI' and 'IDF'; indicate that calculations are based on data sourced from the Electronic Invoices and from the Income Declaration Forms respectively.

5.2 Eliminating markup dispersion across buyers

Table 12 presents results from our counterfactual exercises. The first row in Table 12 reports results under our baseline calibration and definition of products, for different values of the elasticity of substitution ρ. When $\rho = 3$ the change in welfare (or aggregate productivity) is $\hat{C} - 1 = 0.1\%$, as reported in Column 2 To put this magnitude into perspective, note that we are only changing markups for 5% of the total sales. The change in aggregate productivity normalized by the fraction of sales of products with changed markups is 2.3%, as reported in Column 3. When $\rho = 6$, the normalized change in welfare of 7.1%, as reported in Column 5.

The second row reports results under the EAN definition of products. Since only a small fraction of products use EAN codes, the fraction of shocked sales in this exercise

while Edmond et al. (2023) consider elasticities between 5.66 to 29.1.

drops to 2%. The normalized change in welfare in this exercise is larger than in the baseline, 2.9% for $\rho = 3$ and 8.7% for $\rho = 6$. This reflects that EAN products, on average, have a higher number of buyers and more dispersed prices across these buyers compared to the baseline products.

Row 3 reports results using our baseline definition of products but restricting the sample to sellers and buyers located within the city of Santiago. As discussed in Section 2, the assumption that the marginal cost of production is independent of the buyer's identity is more likely to hold when the seller and the buyer are within the same city, as this minimizes potential differences in delivery costs across buyers. The resulting changes in normalized welfare are similar to those our baseline sample when we set $\rho = 3$ (2% vs. 2.3%), while the difference is magnified if we set $\rho = 6$ (5.4% vs. 7%).

The last two rows show results for alternative values of the elasticities in the production functions, σ and η. In Row 4, we we set the elasticity of substitution between sectors $s \in S$ and $s \notin S$ to $\eta = 1$. Lowering this elasticity leads to a somewhat larger change in normalized welfare (2.4% vs. 2.3% when $\rho = 3$). In Row 5 we lower the elasticity of substitution between factors and materials to $\sigma = 0.5$. This parameterization results in a smaller change in aggregate productivity compared to the baseline (1.9% vs. 2.3% when $\rho = 3$). Overall, our baseline results are not very sensitivity to plausible changes in these two elasticities.

5.2.1 Eliminating markup dispersion across products and buyers

We now compare the gains from eliminating markup dispersion across buyers with the gains from eliminating dispersion in average markups across products. In particular, for products $i \in S$ with multiple buyers, we set the counterfactual markup equal to the economy-wide ratio of revenues to costs. For these products, the change in markup is

$$\hat{\mu}_{ii'} = \frac{\mathcal{M}_i}{\mu_{ii'}} \times \frac{\mathcal{M}}{\mathcal{M}_i}.$$

The first ratio in the right hand side is the change from setting the markup $\mu_{ii'}$ equal to the average markup of product i across buyers, $\mathcal{M}_i$, exactly as in our first counterfactual. We obtain this ratio from the invoice data as described in equation (20). The second ratio is the change from setting the product level markup, $\mathcal{M}_i$, equal to the economy wide markup, $\mathcal{M}$. We set $\mathcal{M}_i$ to match the ratio of revenues to costs for firm $f(i)$ from the income declaration forms . We set $\mathcal{M} \equiv \left[\sum_i \mathcal{M}_i^{-1} s_i^r\right]^{-1}$, where $s_i^r = v s_i^c + [1 - v]\, s_i^x$ is the

Table 12: Welfare gains from eliminating markup dispersion across buyers

	shocked sales / total sales	Low elasticity: $\rho = 3$ productivity (%)	(2)/(1)	High elasticity: $\rho = 6$ productivity (%)	(4)/(1)
	(1)	(2)	(3)	(4)	(5)
Baseline	0.05	0.12	2.25	0.37	7.05
EAN	0.02	0.04	2.88	0.13	8.74
Santiago	0.06	0.11	2.01	0.31	5.41
$\eta = 1$	0.05	0.13	2.44	0.42	7.99
$\sigma = 0.5$	0.05	0.10	1.93	0.32	6.19

Notes: This table reports the results of our counterfactual that eliminates the dispersion in markups across buyers of a given input. 'Baseline' refers to our baseline calibration. 'EAN' refers to a network that comprises the EAN products all the buyers of EAN products that are in the Manufacturing and the Retail and Wholesale sectors. 'Santiago' uses our baseline calibration but limits the network to products and firms located in Santiago. '$\eta = 1$' divides our sample of products into two sectors $i \in S$ and $i \notin S$, and sets $\eta = 1$. The last row follows our baseline calibration but sets $\sigma = 0.5$.

share of product i in aggregate sales.

We note again that our model implies that the ratio of revenues to cost for each product is equal to $\frac{\alpha_i}{1-v_i}$. To be able to target $\mathcal{M}_i$ directly without changing the distribution of product sales and purchases, we set $\alpha_i = \mathcal{M}_i \times [1 - v_i]$ for this exercise.[33] As our goal is to compare this counterfactual with our baseline exercise, we keep markups for other products and markups to final sales unchanged.

The first row in Table 13 repeats the baseline counterfactual in Table 12 using the new calibration that targets $\mathcal{M}_i$ rather than α_i. The results of the first counterfactual are similar to those using our baseline calibration. The second row shows the results of eliminating the dispersion in markups both across buyers and across products. The normalized aggregate productivity (or welfare) gain is 3.75% when we set $\rho = 3$, and 9.70% when we

[33]Note that in our baseline calibration, $s_i^c = \left[s_i^m \frac{1}{1-v_i} - s_i^x\right] \frac{1-v}{v}$. To keep s_i^c, s_i^x, and s_i^m unchanged, v_i must be unchanged, so in this exercise we set $\alpha_i = \mathcal{M}_i \times [1 - v_i]$ to target $\mathcal{M}_i$. This recalibration only affects products with $\frac{inputs_{f(i)}^{IDF} \times \theta_{f(i)}}{revenue_{f(i)}^{IDF}} > \frac{s_i^m}{s_i^x}$. For products where $\frac{inputs_{f(i)}^{IDF} \times \theta_{f(i)}}{revenue_{f(i)}^{IDF}} < \frac{s_i^m}{s_i^x}$, we have that $\alpha_i = \mathcal{M}_i \times [1 - v_i] = \frac{revenue_{f(i)}^{IDF}}{cost_{f(i)}^{IDF}} \times \frac{inputs_{f(i)}^{IDF} \times \theta_{f(i)}}{revenue_{f(i)}^{IDF}} = \frac{inputs_{f(i)}^{IDF}}{cost_{f(i)}^{IDF}} \times \theta_{f(i)}$, which coincides with the calibration of α_i described in the previous section.

set $\rho = 6$. These gains are almost twice as large as the gains from eliminating the dispersion across buyers alone. We conclude from this exercise that the aggregate efficiency gains from eliminating markup dispersion across buyers are almost as large as those from eliminating markup dispersion across products.

Table 13: Welfare gains from eliminating markup dispersion across buyers and products

	shocked sales / total sales	Low elasticity: $\rho = 3$		High elasticity: $\rho = 6$	
		productivity (%)	(2)/(1)	productivity (%)	(4)/(1)
	(1)	(2)	(3)	(4)	(5)
Counterfactual 1*	0.05	0.12	2.10	0.32	5.79
Counterfactual 2	0.05	0.21	3.75	0.53	9.70

Notes: 'Counterfactual 1*' reports results from the counterfactual that eliminates the dispersion in markups across buyers of a given input, but using a calibration that directly targets $\mathcal{M}_i$ rather than α_i. 'Counterfactual 2' reports results from the counterfactual that eliminates the dispersion in markups across buyers of a given input and also across inputs under that same calibration.

6 Conclusion

Using a comprehensive dataset of electronic invoices issued by Chilean firms, we document pervasive dispersion in prices that different buyers pay for intermediate goods. While prices are negatively correlated with different measures of buyer and transaction size, observable characteristics of products and of buyer-seller pairs explain a small fraction of the variance of price gaps in the data. The observed price gaps do not appear to reflect differences in the terms of payment or in the cost of supplying different buyers. Using a workhorse model of production networks that reads differences in buyer-specific markups from observed price dispersion, we show that the efficiency gains from eliminating markup dispersion across buyers are comparable in magnitude to those of eliminating markup dispersion across products.

References

Alvarez, Fernando, Martin Beraja, Martin Gonzalez-Rozada, and Pablo Andrés Neumeyer, "From hyperinflation to stable prices: Argentina's evidence on menu cost models," *The Quarterly Journal of Economics*, 2019, *134* (1), 451–505.

Alviarez, Vanessa I, Michele Fioretti, Ken Kikkawa, and Monica Morlacco, "Two-Sided Market Power in Firm-to-Firm Trade," Working Paper 31253, National Bureau of Economic Research May 2023.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium," *The Quarterly Journal of Economics*, 2020, *135* (1), 105–163.

Bornstein, Gideon and Alessandra Peter, "Nonlinear Pricing and Misallocation," *Working paper*, February 2023.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong, "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications," *American Economic Review*, March 2015, *105* (3), 993–1029.

Daruich, Diego and Julian Kozlowski, "Macroeconomic Implications of Uniform Pricing," *American Economic Journal: Macroeconomics*, July 2023, *15* (3), 64–108.

DellaVigna, Stefano and Matthew Gentzkow, "Uniform Pricing in U.S. Retail Chains," *The Quarterly Journal of Economics*, 2019, *134* (4), 2011–2084.

Dhyne, Emmanuel, Amil Petrin, Valerie Smeets, and Frederic Warzynski, "Theory for Extending Single-Product Production Function Estimation to Multi-Product Settings," Working Paper 30784, National Bureau of Economic Research December 2022.

_, Ayumu Ken Kikkawa, and Glenn Magerman, "Imperfect Competition in Firm-to-Firm Trade," *Journal of the European Economic Association*, 2022, *20* (5), 1933–1970.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu, "How Costly Are Markups?," *Journal of Political Economy*, 2023, *131* (7), 1619–1675.

Fontaine, François, Julien Martin, and Isabelle Mejean, "Price discrimination within and across EMU markets: Evidence from French exporters," *Journal of International Economics*, 2020, *124* (C).

Grennan, Matthew, "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, February 2013, *103* (1), 145–77.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Ignatenko, Anna, "Countervailing Power of Firms in International Trade," Working Paper January 2023.

Kaplan, Greg and Guido Menzio, "The morphology of price dispersion," *International Economic Review*, 2015, *56* (4), 1165–1206.

_ , _ , Leena Rudanko, and Nicholas Trachter, "Relative Price Dispersion: Evidence and Theory," *American Economic Journal: Microeconomics*, August 2019, *11* (3), 68–124.

Koike-Mori, Yasutaka and Antonio Martner, "Aggregating Distortions in Networks with Multi-Product Firms," *Working paper*, August 2024.

Loecker, Jan De, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik, "Prices, markups, and trade reform," *Econometrica*, 2016, *84* (2), 445–510.

Macedoni, Luca and Elena Mattana, "Pricing in Firm-to-Firm Trade. Evidence from a Danish Multinational," *Review of World Economics*, July 2023.

Marshall, Guillermo, "Search and Wholesale Price Discrimination," *RAND Journal of Economics*, June 2020, *51* (2), 346–374.

Nakamura, Emi, Jón Steinsson, Patrick Sun, and Daniel Villar, "The elusive costs of inflation: Price dispersion during the US great inflation," *The Quarterly Journal of Economics*, 2018, *133* (4), 1933–1980.

Osotimehin, Sophie and Latchezar Popov, "Misallocation and intersectoral linkages," *Review of Economic Dynamics*, 2023.

Pellegrino, Bruno, "Product Differentiation and Oligopoly: a Network Approach," *Working paper*, February 2023.

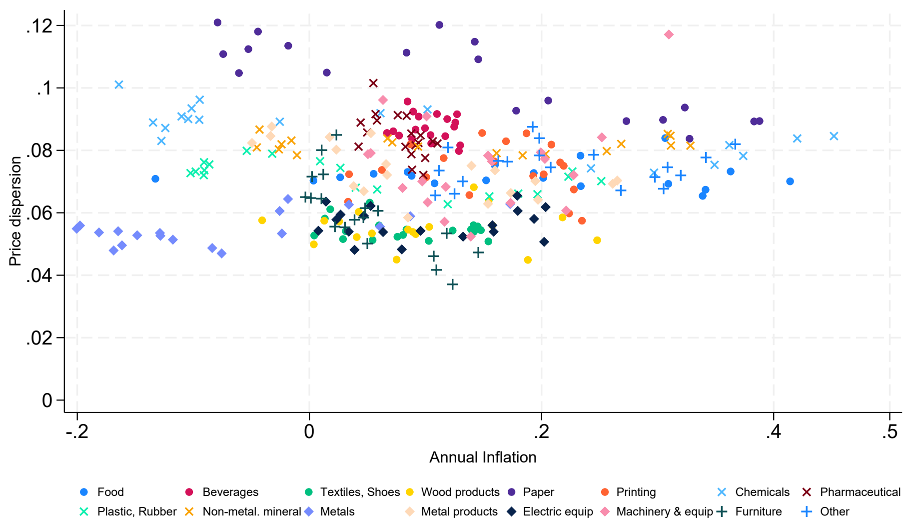
Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, September 2020, *88* (5), 2037–2073.

Sangani, Kunal, "Markups Across the Income Distribution: Measurement and Implications," Working Paper April 2024.

Appendix

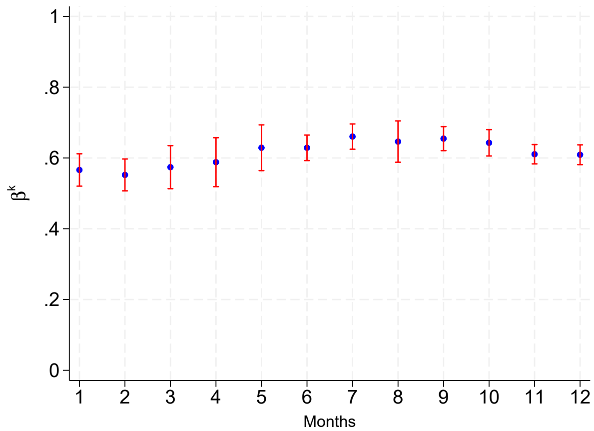
A Additional tables and figures

Figure A.1: Price dispersion and inflation across sectors

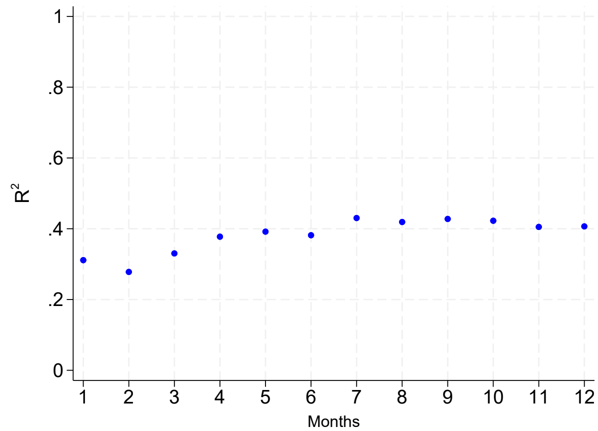


Notes: Each marker represents a sector-month. The x-axis plots the year-to-year change in producer prices for the sector-month. The y-axis plots the price dispersion for the sector-month.

Figure A.2: Persistence in price gaps conditional on a price change



(a) Estimated coefficients β_k



(b) Partial R-squared

Notes: The left panel plots the coefficients β^k from equation (5) for a subset of product-buyer pairs that exhibit non-zero price changes between t and $t-k$, estimated by OLS, and for $k = 1, ..., 12$. The regression includes product fixed effects, and observations are weighted by sales. The red lines represent 95 % confidence intervals. The right panel plots the corresponding partial R-squared of the regressions.

Table A.1: Price dispersion by sector

σ_i

	Food	Beverages	Textiles, Shoes	Wood prod-ucts	Paper	Printing	Chemicals	Pharma-ceutical	Plastic, Rubber	Non-metallic minerals	Metals	Metal prod-ucts	Electric equip	Machinery and equip	Furniture	Other
P1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P5	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00
P10	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.02	0.00
P25	0.03	0.04	0.02	0.02	0.02	0.03	0.03	0.02	0.02	0.05	0.03	0.02	0.02	0.01	0.03	0.02
P50	0.06	0.08	0.04	0.04	0.11	0.05	0.07	0.06	0.06	0.09	0.04	0.06	0.04	0.04	0.07	0.06
P75	0.10	0.12	0.07	0.06	0.16	0.09	0.14	0.12	0.12	0.10	0.07	0.10	0.08	0.09	0.13	0.10
P90	0.14	0.16	0.11	0.09	0.21	0.17	0.21	0.25	0.18	0.14	0.11	0.16	0.13	0.22	0.16	0.13
P95	0.18	0.21	0.14	0.12	0.25	0.24	0.26	0.31	0.22	0.17	0.15	0.20	0.14	0.34	0.21	0.20
P99	0.33	0.28	0.27	0.29	0.34	0.38	0.36	0.45	0.31	0.27	0.20	0.31	0.26	0.45	0.35	0.45
Mean	0.07	0.08	0.05	0.05	0.10	0.08	0.09	0.09	0.08	0.08	0.05	0.08	0.06	0.08	0.08	0.08
Share multibuyer	0.53	0.80	0.55	0.55	0.49	0.37	0.58	0.74	0.38	0.67	0.77	0.37	0.65	0.16	0.58	0.31

Notes: The table reports the distribution of the standard deviation of prices across buyers of the same product within the different manufacturing subsectors in our baseline sample. The distributions are based on the subsample or products that were purchased by more than one buyer in June 2023. 'Share multibuyer' refers to the share of subsector's sales that is accounted for by products that ere purchased by more than one buyer.

Table A.2: Persistence in price gaps

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$	$k=11$	$k=12$
β_k	0.860***	0.824***	0.774***	0.732***	0.750***	0.712***	0.698***	0.686***	0.6878***	0.664***	0.619***	0.620***
	(0.0102)	(0.0113)	(0.0132)	(0.01695)	(0.0145)	(0.0145)	(0.0168)	(0.0225)	(0.0182)	(0.0162)	(0.0147)	(0.0151)
Partial R2	0.720	0.646	0.611	0.568	0.562	0.513	0.498	0.486	0.467	0.460	0.425	0.428
Observations	1,805,384	1,675,076	1,768,055	1,489,223	1,475,565	1,460,801	1,447,728	1,395,106	1,3110,79	1,251,503	1,184,824	1,085,192

Notes: The table reports the results depicted in Figure 2, which correspond to estimating equation (5) by OLS separately for $k = 1, ..., 12$. All columns include product-level fixed effects.

Table A.3: Summary statistics on network

Number of products	240,763
Number of firms	123,111
Possible product-buyer pairs	> 30 billion
Product-buyer pairs with positive sales	> 3 million
Sales of $i \in S$ / total sales	0.30
Intermediate sales of $i \in S$ / total sales	0.9
Intermediate sales of i with multiple buyers / total sales	0.052

Notes: The table summarizes the network of products and firms used for our baseline counterfactual. $i \in S$ refers to products for which product codes are available.

B Data

B.1 Additional datasets

Here we outline additional datasets we use in the paper:

Form 22 (F22 or Income Declaration Form): This is a form submitted by firms to compute their annual tax payments, and therefore contains total sales in the year, intermediate inputs purchases, salaries, among other variables. We use this dataset for regressions where we measure buyer's size with their total sales, and for calibration of the model.

Affidavit 1887 (DJ1887): This an annual form submitted by firms, where they declare all their employees with either a full-time or a part-time job contract. For each firm, it contains what months each worker was employed, her net salaries and whether it was a full-time or a part-time job.

Affidavit 1879 (DJ1879): This an annual form submitted by firms, where they declare all their contracted or short-term workers. For each firm, it contains how many months each worker was employed, and how much it was retained for taxes and social security.

We use DJ1887 and DJ1879 for two purposes in the paper. First, our baseline sample uses firms (either as buyers or sellers) that have at least one paid worker in 2021. Second, for regressions that use a measure of employment at the firm level, we sum across all workers$\times$months in both forms and divide them by 12.

B.2 Data cleaning

We take the following steps to organize and clean the transaction data in EI:

i. We drop any transaction whose value is above 100 billion CLP (10^10), and price or quantity is non-negative.

ii. We discard invoices where the seller and buyer ID are the same

iii. We eliminate invoices that are cancelled or deemed invalid. For this, we use credit notes which give *new* or *updated* total amounts for each invoice. We drop documents where the new total amount differs by more than 1% with respect to the original total amount in the invoice

iv. We keep firms that in 2021 had at least one paid worker, and a valid F22.

v. Then, we define a product using the seller ID, the product code and the municipality for the seller declared in an invoice. Firms can report up to five codes, and whenever they declare the code is EAN we use it. Firms _write_ the municipality, thus it is open to typos. When we cannot identify a selling municipality, we define the product only using the seller ID and product code.

vi. To further refine our product definition, we only consider codes that fulfill the following: (i) are at least 3 characters long, (ii) have a number, (iii) for a given text description of the product, there is a unique product code, and (iv) for a given product code, there is a unique text description of the product.

vii. We eliminate transactions where the product's recorded price deviates from the average price recorded in other transactions from the same month by more than 1 log point.

C Model derivations

C.1 Equilibrium characterization

The cost of the input bundle for product i is given by

$$mc_i = \left[[1 - \bar{\alpha}_i] \, w^{1-\sigma} + \bar{\alpha}_i q_i^{1-\sigma} \right]^{\frac{1}{1-\sigma}} . \tag{C.1}$$

Here the cost of the fixed factors, and q_i is the cost of the intermediate input bundle:

$$q_i = \left[\sum_j \bar{\omega}_i^j \left[q_i^j \right]^{1-\eta} \right]^{\frac{1}{1-\eta}} , \tag{C.2}$$

with

$$q_i^j = \left[\sum_{i' \in j} \bar{\omega}_{i'i} p_{i'i}^{1-\rho} \right]^{\frac{1}{1-\rho}} . \tag{C.3}$$

The price of the final good is:

$$P = \left[\sum_j \gamma^j \left[P^j \right]^{1-\eta} \right]^{\frac{1}{1-\eta}}, \tag{C.4}$$

with

$$P^j = \left[\sum_{i \in j} \bar{\gamma}_i^j p_{ic}^{1-\rho} \right]^{\frac{1}{1-\rho}}. \tag{C.5}$$

Demands are given by:

$$x_{ii'} = \bar{\omega}_{i'i} \left[\frac{p_{ii'}}{q_{i'}^j} \right]^{-\rho} m_{i'}^j; \quad c_i = \bar{\gamma}_i \left[\frac{p_{ic}}{P^{j(i)}} \right]^{-\rho} C^{j(i)}. \tag{C.6}$$

with

$$m_i^j = \bar{\omega}_i^j \left[\frac{q_i^j}{q_i} \right]^{-\eta} m_i,$$
$$m_i = \bar{\alpha}_i \left[\frac{q_i}{mc_i} \right]^{-\sigma} b_i, \tag{C.7}$$

and

$$C^j = \bar{\gamma}^j \left[\frac{P^j}{P} \right]^{-\eta} C. \tag{C.8}$$

Demand for the fixed factor is

$$l_i = [1 - \bar{\alpha}_i] \left[\frac{w}{mc_i} \right]^{\sigma} y_i. \tag{C.9}$$

An equilibrium for this economy is given by product level prices, $\{p_{ii}\}$, $\{p_{ic}\}$, costs $\{mc_i, q_i\}$, $\left\{q_i^j\right\}$, quantities $\{c_i\}$, $\{y_i\}$ and inputs choices: $\{l_i\}$, $\{m_i\}$, $\left\{m_i^j\right\}$, and $\{x_{ii'}\}$, final prices $\left\{P^l\right\}$ P and bundles$\left\{C^j\right\}$, C and factor prices:. such that: $w = 1$ and the pricing equations (11), (C.1), (C.2), (C.3), (C.4), (C.5), the demands (C.6), (C.7), (C.8), (C.9), and market clearing (9), (10) are satisfied.

C.2 Derivation of Equation (18)

This Section derives (18) in the text. Equations (16) and (17) imply

$$1=\sum_i s_i^l \widehat{mc}_i^{\sigma}\left[\frac{c_i}{y_i}\hat{p}_{ic}^{-\rho}\left[\hat{P}^j\right]^{\rho-\eta}\hat{P}^{\eta}\hat{C}+\left[1-\frac{c_i}{y_i}\right]\sum_{i'} s_{ii'}\hat{p}_{ii'}^{-\rho}\left[\hat{q}_{i'}^{j}\right]^{\rho-\eta}\hat{q}_{i'}^{\eta-\sigma}\widehat{mc}_{i'}^{\sigma}\hat{y}_{i'}\right]$$

Under the assumptions that $\alpha_i=0$ and $c_i/y_i=0$ for $i\in N_u$ and $\alpha_i=1$ and $c_i/y_i=1$ for $d\in N_d$, this simplifies to

$$1=\sum_{i\in N_u} s_i^l\left[\sum_{i'\in N_d} s_{ii'}\hat{p}_{ii'}^{-\rho}\hat{P}^{\rho}\hat{C}\right],$$

Substituting for $\hat{P}$ and $\hat{p}_{ic}$ and rearranging yields

$$\hat{C}=\left[\sum_{i\in N_u} s_i^l\left[\sum_{i'\in N_d} s_{ii'}\hat{p}_{ii'}^{-\rho}\right]\right]^{-1}\left[\sum_{i\in N_d} s_i^c \hat{mc}_i^{1-\rho}\right]^{\frac{-\rho}{1-\rho}}.$$

Finally, substituting $\hat{p}_{ii'}$ for $i\in N_u$ and for $\hat{mc}_i$ for $i\in N_d$ we obtain

$$\hat{C}=\left[\sum_{i\in N_u} s_i^l\sum_{i'\in N_d} s_{ii'}\hat{\mu}_{ii'}^{-\rho}\right]^{-1}\left[\sum_{i\in N_d} s_i^c\sum_{i'\in N_u}\omega_{i'i}\hat{\mu}_{i'i}^{1-\rho}\right]^{\frac{\rho}{\rho-1}}.$$

Note that for $i\in N_u$, $s_i^l=\frac{wl_i}{wL}=\frac{\overline{\mathcal{M}_u}}{\mathcal{M}_i}\frac{\sum_{i'\in N_d}p_{ii'}x_{ii'}}{\sum_{i\in N_u}\sum_{i'\in N_d}p_{ii'}x_{ii'}}=\frac{\overline{\mathcal{M}_u}}{\mathcal{M}_i}s_i^x$, where $\overline{\mathcal{M}_u}\equiv\frac{\sum_{i\in N_u}\sum_{i'\in N_d}p_{ii'}x_{ii'}}{wL}=\left[\sum_{i\in N_u}\frac{1}{\mathcal{M}_i}s_i^x\right]^{-1}$. For $i\in N_d$:

$$\begin{aligned}
s_i^c\omega_{i'i}\hat{\mu}_{i'i}&=\frac{p_{ic}c_i}{\sum_{i\in N_d}p_{ic}c_i}\frac{p_{i'i}x_{i'i}}{\sum_{i'\in N_u}p_{i'i}x_{i'i}}\frac{\mathcal{M}_{i'}}{\mu_{i'i}}\\
&=\frac{p_{ic}c_i}{\sum_{i'\in N_u}p_{i'i}x_{i'i}}\frac{\sum_{i'\in N_u}\sum_{i\in N_d}p_{i'i}x_{i'i}}{\sum_{i\in N_d}p_{ic}c_i}\frac{\sum_{i\in N_d}p_{i'i}x_{i'i}}{\sum_{i'\in N_u}\sum_{i\in N_d}p_{i'i}x_{i'i}}\frac{p_{i'i}x_{i'i}}{\sum_{i\in N_d}p_{i'i}x_{i'i}}\frac{\mathcal{M}_{i'}}{\mu_{i'i}}\\
&=\frac{\mathcal{M}_i}{\overline{\mathcal{M}_d}}s_{i'}^x s_{i'i},
\end{aligned}$$

where $\overline{\mathcal{M}_d}\equiv\frac{\sum_{i\in N_d}p_{ic}c_i}{\sum_{i\in N_d}\sum_{i'\in N_u}p_{i'i}x_{i'i}}=\left[\sum_{i\in N_d}\frac{1}{\mathcal{M}_i}s_i^c\right]^{-1}$. Substituting above and rearranging yields equation (18) in the text.

Finally, using the notation $\xi_i\equiv\sum_{i'\in N_d}\hat{\mu}_{ii'}^{-\rho}s_{ii'}$ to summarize the cross-buyer dispersion in markups for product i, and assuming that ξ_i is uncorrelated with $\mathcal{M}_i$, we can write the first bracket in equation (18) as

$$\sum_{i \in N_u} \frac{\overline{\mathcal{M}_u}}{\mathcal{M}_i} \xi_i s_i^x = \sum_{i \in N_u} \frac{\overline{\mathcal{M}_u}}{\mathcal{M}_i} s_i^x \times \sum_{i \in N_u} \xi_i s_i^x = \sum_{i \in N_u} \xi_i s_i^x. \tag{C.10}$$

In addition, if the wedges are uncorrelated with the average markups set by the buyers, $\sum_{i' \in N_d} \mathcal{M}_{i'} \hat{\mu}_{ii'}^{-\rho} s_{ii'} = \sum_{i' \in N_d} \mathcal{M}_{i'} s_{ii'} \xi_i$, we can write the second bracket in (18) as

$$\sum_{i \in N_u} s_i^x \sum_{i' \in N_d} \frac{\mathcal{M}_{i'}}{\overline{\overline{\mathcal{M}_d}}} \hat{\mu}_{ii'}^{-\rho} s_{ii'} = \sum_{i \in N_u} \left[\xi_i \sum_{i' \in N_d} \frac{\mathcal{M}_{i'}}{\overline{\overline{\mathcal{M}_d}}} s_{ii'} \right] s_i^x. \tag{C.11}$$

Finally, assuming that $\sum_{i \in N_u} \left[\xi_i \sum_{i' \in N_d} \mathcal{M}_{i'} s_{ii'} \right] s_i^x = \left[\sum_{i \in N_u} \xi_i s_i^x \right] \times \left[\sum_{i \in N_u} \sum_{i' \in N_d} \mathcal{M}_{i'} s_{ii'} s_i^x s_i^x \right] = \left[\sum_{i \in N_u} \xi_i s_i^x \right] \overline{\mathcal{M}_d}$, and substituting (C.10) and (C.11) into (18) we obtain expression (19) in the text.

D Calibration details

D.1 Apportioning firm-level inputs to products

Here we show how to apportion firm-level inputs to products under our assumptions in Section 5. First, note that if firm's uses the same input bundle across its multiple products, then the share of input i' in the production of i is equal to the share of input i in firm's $f(i)$ total cost:

$$\frac{p_{i'i}x_{i'i}}{\sum_{i\in f(i)} p_{i'i}x_{i'i}} = \frac{\sum_{i'} p_{i'i}x_{i'i}}{\sum_{i\in f(i)}\sum_{i'} p_{i'i}x_{i'i}} = \frac{wl_i + \sum_{i'} p_{i'i}x_{i'i}}{\sum_{i\in f(i)}\left[wl_i + \sum_{i'} p_{i'i}x_{i'i}\right]}$$

If in addition the firm sets the same average markup across all its products, then this share is equal to the share of product i in firm's $f(i)$ sales.

$$\frac{wl_i + \sum_{i'} p_{i'i}x_{i'i}}{\sum_{i\in f(i)}\left[wl_i + \sum_{i'} p_{i'i}x_{i'i}\right]} = \frac{p_i^c c_i + \sum_{i'} p_{ii}x_{ii'}}{\sum_{i\in f(i)}\left[p_i^c c_i + \sum_{i'} p_{ii}x_{ii'}\right]}.$$

Finally, if the share of sales that go to final consumers is the same for all products of the same firm, we obtain

$$\frac{p_i^c c_i + \sum_{i'} p_{ii}x_{ii'}}{\sum_{i\in f(i)}\left[p_i^c c_i + \sum_{i'} p_{ii'}x_{ii'}\right]} = \frac{\sum_{i'} p_{ii}x_{ii'}}{\sum_{i\in f(i)}\sum_{i'} p_{ii}x_{ii'}} \equiv \kappa_i.$$

Thus, $\frac{p_{i'i}x_{i'i}}{\sum_{i\in f(i)} p_{i'i}x_{i'i}} = \frac{\sum_{i'} p_{i'i}x_{i'i}}{\sum_{i'}\sum_{i\in f(i)} p_{i'i}x_{i'i}} = \kappa_i$. We can then obtain the equations in the text:

$$s_{ii'} \equiv \frac{x_{ii'}}{\sum_b x_{ib}} = \frac{x_{if(i')}}{\sum_b x_{ib}} \times \frac{x_{ii'}}{x_{if(i')}} = \frac{x_{if(i')}}{\sum_{i'} x_{ii'}}\kappa_i.$$

and

$$s_i^m = \frac{\sum_{i'} p_{i'i}x_{i'i}}{\sum_{i'}\sum_b p_{i'b}x_{i'b}} = \frac{\sum_{i'}\sum_{i\in f(i)} p_{i'i}x_{i'i}}{\sum_{i'}\sum_b p_{i'b}x_{i'b}} \times \frac{\sum_{i'} p_{i'i}x_{i'i}}{\sum_{i'}\sum_{i\in f(i)} p_{i'i}x_{i'i}} = s_{f(i)}^m\kappa_i.$$

D.2 Calibrating value added shares

As described in the text, an input in the calibration is the ratio of value added to revenues for each product i, v_i. This ratio, which can be obtained from the income declaration forms, must satisfy (for positive prices and quantities) $1 - v_i = \frac{\sum_{i'} p_{i'i}x_{i'i}}{p_{ic}c_i + \sum_{i'} p_{ii'}x_{ii'}} \leq \frac{\sum_{i'} p_{i'i}x_{i'i}}{\sum_{i'} p_{ii'}x_{ii'}} = \frac{s_i^m}{s_i^x}$. We thus compute the ratio of input purchases to revenues in the income declaration forms and set $1 - v_i = \min\left\{\frac{inputs_{f(i)}^{IDF} \times \theta_{f(i)}}{revenue_{f(i)}^{IDF}}, 1 - \frac{s_i^m}{s_i^x}\right\}$.

E Solution Algorithm

We now present the solution algorithm, the set of equations, and the notation for Matlab. We evaluate the response of the economy to a change in intermediate markups, $\hat{\mu}_{if}$, and assume that the change in markups to final consumers are the same for all products of the same firm $\hat{\mu}_{ic} = \hat{\mu}_{f(i)c}$,

i. **Step 1:** Solve for all price changes. We proceed in the following steps.

   (a) **Make a guess for the change in marginal cost of each firm:** $\widehat{mc}^*_f \equiv \widehat{mc}_f^{1-\rho}$. This is an $NF_T \times 1$ vector. We iterate on $\widehat{mc}_f^{1-\rho}$instead of $\widehat{mc}_f$ to speed up code.

   (b) **Define the firm to firm weighted markups:** $\hat{\mu}^*_{ff'} \equiv \sum_{j \in f} \theta_{jf'} \hat{\mu}_{jf'}^{1-\rho}$.

   (c) **Iterate the following system to find** $\widehat{mc}^*_f$

$$\hat{p}^*_{ff'} \equiv \hat{\mu}^*_{ff'} \times \widehat{mc}^*_f,$$

and

$$\hat{q}^{M*}_{f'} = \left[\sum_{f \in M} \hat{p}^*_{ff'} \right]^{\frac{1-\eta}{1-\rho}} ; \hat{q}^{O*}_{f'} = \left[\sum_{f \in O} \hat{p}^*_{ff'} \right]^{\frac{1-\eta}{1-\rho}} .$$

and

$$\hat{q}^*_f = \nu^M_f \hat{q}^{M*}_f + \left[1 - \nu^M_f \right] \hat{q}^{O*}_f .$$

and

$$\widehat{mc}^*_f = \left[1 - \alpha_f + \alpha_f \hat{q}_f^{*\frac{1-\sigma}{1-\eta}} \right]^{\frac{1-\rho}{1-\sigma}} .$$

(d) Compute all prices

$$
\begin{aligned}
\widehat{mc}_f &= \left[\widehat{mc}_f^*\right]^{\frac{1}{1-\rho}} \\
\hat{p}_{if'} &= \hat{\mu}_{if'}\widehat{mc}_{f(i)} \\
\hat{p}_{ic} &= \hat{\mu}_{f(i)c}\widehat{mc}_{f(i)} \\
\hat{q}_{f'}^M &= \left[\hat{q}_{f'}^{M*}\right]^{\frac{1}{1-\eta}} \\
\hat{q}_{f'}^O &= \left[\hat{q}_{f'}^{O*}\right]^{\frac{1}{1-\eta}} \\
\hat{q}_{f'} &= \left[\hat{q}_{f'}^*\right]^{\frac{1}{1-\eta}}
\end{aligned}
$$

$$
\begin{aligned}
\hat{P}^M &= \left[\sum_{f\in M} \gamma_f \hat{p}_{fc}^{1-\rho}\right]^{\frac{1}{1-\rho}} \\
\hat{P}^O &= \left[\sum_{f\in O} \gamma_f \hat{p}_{fc}^{1-\rho}\right]^{\frac{1}{1-\rho}}
\end{aligned}
$$

and

$$
\hat{P} = \left[\beta^M \left[\hat{P}^l\right]^{1-\eta} + \left[1-\beta^M\right]\left[\hat{P}^l\right]^{1-\eta}\right]^{\frac{1}{1-\eta}}.
$$

ii. **Step 2:** Solve for changes in the total quantity of each product relative to the change in final consumption, $\frac{\hat{y}_i}{\hat{C}}$. We proceed in the following steps.

(a) **Compute change in relative final demand of each product:**

$$
\frac{\hat{c}_i}{\hat{C}} = \left[\frac{\hat{p}_{ic}}{\hat{P}^{k(i)}}\right]^{-\rho}\left[\frac{\hat{P}^{k(i)}}{\hat{P}}\right]^{-\eta}.
$$

(b) **For products that are only sold to final consumers, set**

$$
\frac{\hat{y}_i}{\hat{C}} = \frac{\hat{c}_i}{\hat{C}}.
$$

(c) **Then guess $\frac{\hat{y}_i}{\hat{C}}$ for products $[1 : NP_{XT}]$, and compute**

$$
\frac{\hat{b}_f}{\hat{C}} = \sum_{i\in f} \frac{b_i}{b_f}\frac{\hat{y}_i}{\hat{C}}.
$$

and update the guess using

$$\frac{\hat{y}_i}{\hat{C}} = \widehat{FDEM}_i + \sum_f INTDEM_{if} \frac{\hat{b}_f}{\hat{C}}.$$

with $\widehat{FDEM}_i \equiv \frac{c_i}{y_i} \frac{\hat{c}_i}{\hat{C}}$, and $\boldsymbol{INTDEM_{if}} = \left[1 - \frac{c_i}{y_i}\right] \frac{x_{if}}{\sum_f x_{if}} \left[\frac{\hat{p}_{if}}{\hat{q}_f^{k(i)}}\right]^{-\rho} \left[\frac{q_f^{k(i)}}{\hat{q}_f}\right]^{-\eta} \left[\frac{\hat{q}_f}{\widehat{mc}_f}\right]^{-\sigma}$

(d) **From the labor market clearing, solve:**

$$\hat{C} = \left[\sum_i \left[\frac{1}{\widehat{mc}_f}\right]^{-\sigma} \frac{\hat{b}_f}{\hat{C}} \frac{l_f}{L}\right]^{-1}.$$