Aggregating Distortions in Networks with Multi-Product Firms

Yasutaka Koike-Mori *
UCLA

Antonio Martner
UCLA & Central Bank of Chile

Job Market Paper

[version: October 20, 2024]

Click here for the latest version.

Abstract

We investigate the role of multiproduct firms in shaping resource misallocation within production networks and its impact on aggregate total factor productivity (TFP) growth. Using administrative data on product transactions between all the formal Chilean firms, we provide evidence that demand shocks to one product affect the production of other products within the same firm, suggesting that firms engage in joint production. We develop a framework to measure resource misallocation in production networks with joint production, deriving non-parametric sufficient statistics to quantify these effects. Applying the framework to Chile, we find that reallocation effects, considering joint production, explain 86% of the observed aggregate TFP growth for the 2016-2022 period. Ignoring joint production leads to overestimating resource misallocation.

*Yasutaka Koike-Mori is deeply indebted to his doctoral committee members David Baqaee (co-chair), Ariel Burstein (co-chair), Hugo Hopenhayn, Jonathan Vogel, Oleg Itskhoki, Andy Atkeson for their continuous support. Pablo Fajgelbaum, Michael Rubens, Colin Hottman, Aaron Flaaen, Xiang Ding, and Kosuke Aoki provided valuable comments. Yasutaka Koike-Mori acknowledges the generous support the Research and Statistics Department at the Federal Reserve Board (FRB) provided during his dissertation fellowship. We also thank seminar participants at Penn State University, the Third Global Economic Network Conference, the GSE-OSIPP-ISER Joint Conference, Keio Univesity, Waseda University, Hitotsubashi University, the Midwest Macroeconomics Meeting, the Federal Reserve Board, and the Central Bank of Chile.

1 Introduction

Resource misallocation is known to be an important driver of aggregate Total Factor Productivity (TFP) growth. One of the lessons from the literature on misallocation has been the importance of input-output networks in quantifying the extent and importance of misallocation. [1] However, previous research has typically assumed that firms produce a single product.

This is potentially a strong assumption, given that most firm-to-firm transactions involve multiproduct firms. For example, using the universe of firm-to-firm transactions from Chile, we find that 75% of firms sell multiple products, that account for 99% of the total firm-to-firm transaction value. The prevalence of multiproduct firms introduces a new channel for reallocation; resources may be inefficiently allocated across products within a firm.

How do multiproduct firms shape resource allocation within general production networks? In this paper we provide an answer and quantify the importance of this channel for understanding misallocation and aggregate TFP growth.

One of the challenges with modeling multiproduct firms is assigning inputs to outputs within the firm. Typically, detailed production data tracking the input usage for each specific product are unavailable. Furthermore, in some cases, such questions may not even be coherent. For example, a dairy farmer produces meat and milk simultaneously; there is no clear way to allocate the inputs between milk and meat production. Similarly, an oil refinery produces diesel and gasoline concurrently, with output ratios dictated by chemistry rather than economics.

This paper tackles this problem by modelling firm's technologies jointly producing all outputs. That is, the firm's technology maps a collection of inputs into a collection of outputs, and there need not be a well-defined notion of individual products production functions. The shape of this technology set how easy it is for a firm to alter its product mix, which shapes the importance of resource misallocation inside the firm.

We first ask whether multiproduct firms can be relabeled to be a collection of independent single-product firms, which is often assumed in the literature (Klette and Kortum (2004); Bernard et al. (2011); Hottman et al. (2016); Mayer et al. (2021)). If this assumption holds, existing aggregation theorems could be applied by treating each product as a separate firm. To test this assumption, we exploit the heterogeneous exposure of different

[1] For example, Baqaee and Farhi (2020) found that by considering production linkages, resource reallocation explains about half of the aggregate TFP growth in the United States from 1997 to 2015.

product-buyer pairs within the same firm to COVID-19 lockdowns in Chile. We show that a demand shock for one product affects the production of other products within the same firm. This finding suggests that firms cannot be relabeled to be a collection of independent single-product firms. Therefore, firms engage in joint production, which implies using common inputs to produce different products. This finding aligns with recent findings by Ding (2023).

We build a theoretical framework to consider multiproduct firms engaged in joint production within production networks and assess the aggregate TFP growth implications. Firms involved in joint production face constraints when adjusting their product mix, which dampen the effects of reallocation. The more difficult it is for a firm to adjust its product mix, the more strongly the multiproduct firm attenuates the reallocation effect.

We derive nonparametric sufficient statistics to characterize allocative efficiency with multiproduct firm that engages in the joint production. We construct these statistics using observed data to decompose measured aggregate TFP growth into allocative efficiency and technology. Our theory assumes that firms engage in joint production, generalizing Baqaee and Farhi (2020)'s production network structure.

Our sufficient statistics detect if each firm involves joint production; some firms may have independent product lines, while others may face strict constraints on adjusting their product mix. Our insight is that we do not need to identify the production structure to understand how joint production has affected observed aggregate TFP growth ex-post; the observed relative price changes of the products within the firm contain information on the transformation ratio between the two goods. This non-parametric detection without specifying the joint production structure comes at the cost of sacrificing the ability to conduct counterfactual analysis.

We implement our framework using a unique dataset that tracks firm-product level transactions and prices for the universe of formal firms in Chile from the Chilean Internal Revenue Service. Tax enforcement in Chile requires that all formal firms declare their invoices with other firms, providing information on all products, quantities, and prices traded between firms. We also access tax accounting declarations, obtaining data on each firm's revenue and input expenditures, including capital and labor.

While our framework can fit any wedge, we assume that product-level output wedges (difference between price and marginal cost) are the sole source of economic inefficiency. We derive the "cumulative markup", which is composed of two elements. The first is the interaction between a firm-product pair network centrality measure and a gross domestic product (GDP) share. We compute both components without any parametric assumption. The second is a product-level markup, which we build using two different strategies. First, we estimate markups using an off-the-shelf production function approach (Dhyne et al. (2022)), where markups are affected by parametric assumptions. Second, we assume product markups equal firm-level average markup and quantify them using the accounting approach (revenue over cost). This approach allows us to eliminate all parametric assumptions. Quantitatively, both approaches generate nearly equivalent aggregate results.

We find that changes in allocative efficiency can explain 86% of TFP growth from 2016 to 2022 in Chile. We also show that ignoring joint production leads to a misleading picture and dramatically overstates the importance of reallocations. We interpret this result as multiproduct firms facing constraints in adjusting their product mix. These constraints limit the scope for product-level productive resource reallocation within each firm, affecting allocative efficiency and aggregate TFP growth.

The allocative efficiency contribution of multiproduct firms engaging in joint production increases during economic disruptions. After the COVID-19 pandemic and the subsequent high inflation period, multiproduct firms engaging in joint production increased their contribution to aggregate TFP growth.

While our paper has primarily focused on nonparametric ex-post analysis based on observed data, we introduce a complementary ex-ante approach in our final analysis. This approach enables counterfactual analysis by imposing parametric structure on the production process. We propose a flexible system that extends the nested-CES framework to incorporate joint production. Using this, we demonstrate how the ”distance to the frontier” - a key measure of resource misallocation in the literature - changes when accounting for joint production, showing how previous theoretical results on misallocation are changed.

Related Literature

This work contributes to and connects different strands of literature. We incorporate multiproduct firms and joint production to extend the rapidly developing literature on misallocation in production networks and growth accounting (Restuccia and Rogerson (2008); Hsieh and Klenow (2009); Baqaee and Farhi (2020); Bigio and La’O (2020); Osotimehin and Popov (2023)). Our theory provides a tool for growth accounting (Solow (1957); Hulten (1978); Basu and Fernald (2002); Petrin and Levinsohn (2012); Baqaee and Farhi (2020);

Baqaee et al. (2023)) that decomposes aggregate TFP growth into technology and allocative efficiency under joint production in networks and generalizes existing methods to consider multiproduct firms.

We contribute to the literature on multiproduct firms (Bernard et al. (2010, 2011); Mayer et al. (2014); Hottman et al. (2016); Mayer et al. (2021)) by showing that they differ from collections of single-product firms in how they transmit shocks due to their joint production activities. Our empirical strategy contrasts with other network studies that examine downstream propagation of supply shocks (Boehm et al. (2019); Carvalho et al. (2020); Fujiy et al. (2022); Bai et al. (2024)). Instead, we focus on how demand shocks to one product affect the production of other products within the same firm. We contrast and discuss our results with those of existing studies that discuss the spillover of demand shocks within the same firm (Giroud and Mueller (2019); Almunia et al. (2021); Ding (2023)).

We extend recent work on joint production (Boehm and Oberfield (2023); Carrillo et al. (2023); Ding (2023)) by revealing the allocative efficiency implications of joint production patterns in networks. [2] These studies focus on how joint production patterns are systematically linked to input proximity. In contrast, our analysis takes these patterns as given and examines their implications for resource allocation and aggregate TFP growth.

Our empirical application uses a comprehensive product-level trade database from Chile to quantify misallocation. This approach contrasts with the prior literature on production networks and misallocation, which uses industry-level input-output tables instead of firm-to-firm data. For example, Baqaee and Farhi (2020) impute US Compustat data using an industry-level input-output table. Even when firm-level transaction data are available, incomplete price information limits the analysis. Furthermore, Kikkawa (2022) examines firm pair-specific markups based on a theoretical model using Belgian interfirm transaction data. Finally, Burstein et al. (2024) uses the same dataset as ours but complements our work by analyzing misallocations arising from different buyers receiving varying prices for the same product.

Lastly, this work is related to the literature on production function estimation. Notably, estimation methods for joint production have been developed recently Dhyne et al. (2017, 2022); De Loecker et al. (2016); Valmari (2023); Cairncross and Morrow (2023). We have not developed any theoretical innovation in this area; however, our application of

[2] The literature on joint production contains seminal works by Powell and Gruen (1968); Diewert (1971); Lau (1972); Hall (1973, 1988). Despite its long-standing nature, the extant literature has been revived due to the recent availability of detailed firm and product-level data

these methods is more comprehensive than previous studies. Unlike previous papers that estimate production functions for a specific industry or subsample of the economy, we apply the method of Dhyne et al. (2022) to estimate multiproduct production functions. Our estimation covers the universe of products traded in Chile by formal firms from 2016 to 2022.

The rest of the paper is organized as follows. Section 2 presents the data and motivating facts, highlighting the dominance of multiproduct firms. It also presents empirical evidence suggesting firms engage in joint production. Section 3 outlines the theoretical framework, deriving the nonparametric sufficient statistics for measuring allocative efficiency explained by multiproduct firms. Section 4 details the data and the construction of sufficient statistics. Section 5 applies the framework to decompose aggregate TFP growth in Chile for the 2016–2022 period. Section 6 presents ex-ante structural results and counterfactual analysis. Finally, Section 7 concludes.

2 Reduced form evidence

This section presents empirical facts that motivate our theory of multiproduct firms engaging in joint production in production networks.

The primary aim of our empirical analysis is to test whether firms can be treated as collections of independent single-product liness. If this were the case, existing aggregation theorems could be applied by considering each product as a separate firm. Our findings, however, reject this separability assumption, revealing interdependencies among products within firms. This result motivates the need for a new aggregation theorem that we develop in subsequent sections.

We use data from the Chilean Internal Revenue Service (SII), covering all formal firms in Chile.[3] The cross-sectional analysis uses 2018, a year unaffected by any major shock. We then employ monthly data from January 2019 to December 2021 to test for joint production. We exploit the unexpected nature of early COVID-19 lockdowns as a source of

[3]This study was developed within the scope of the research agenda conducted by the Central Bank of Chile (CBC) in economic and financial affairs of its competence. The CBC has access to anonymized information from various public and private entities by virtue of collaboration agreements signed with these institutions. To secure the privacy of workers and firms, the CBC mandates that the development, extraction, and publication of the results should not allow the identification, directly or indirectly, of natural or legal persons. Officials of the CBC processed the disaggregated data. All the analysis was implemented by the authors and did not involve nor compromise the Chilean IRS. The information contained in the databases of the Chilean IRS is of a tax nature originating in self-declarations of taxpayers presented to the Service; therefore, the veracity of the data is not the responsibility of the Service.

exogenous variation in product-specific demands.

The SII provides detailed information on firm-to-firm transactions through electronic tax documents. This dataset captures every product, quantity, and price traded between formal Chilean firms, containing data on over 15 million unique firm-specific product descriptions.[4] For 2018, we sum the real-time quantity traded and the value for every buyer firm–seller firm–product triplet transaction. We divide the value traded over its quantity to obtain average yearly prices for every triplet. We use the 2018 data to describe the main features of the firm-to-firm trade patterns.

Fact 1: Multiproduct Firms Dominate Domestic Intermediate Inputs Trade

Of all firms, 75% produce multiple products, and these firms account for 98.94% of intermediate input transaction value. Table 1 illustrates the distribution of products per firm, weighted by firm-to-firm transaction values.

[4]The specific invoice variable is called "detail", which is inherently firm-specific and can differ between firms even for the same product. For example, one supermarket might declare selling "Sprite can 330cc" while another declares selling "Sprite 330". This variation across sellers does not affect our analysis in this Section as we do not compare identical products across firms.

Table 1: Distribution of Product Numbers

Percentile	Number of products (Unweighted)	Number of products (Weighted by transaction value)
1%	1	1
5%	1	2
10%	1	4
25%	2	36
50%	7	475
75%	26	2,459
90%	119	32,195
95%	290	37,422
99%	1,253	62,372

Notes: The Table presents the distribution of product numbers for 2018. The left column shows the number of products without weighting, while the right column displays the number of products weighted by intermediate product transaction volumes of the firms.

Fact 2: Demand Shocks to One Product Affect the Production of Other Products within the Same Firm

Given the prevalence of multiproduct firms established in Fact 1, we examine the role of multiproduct firms in production networks. Most of the existing work (Bernard et al. (2010, 2011); Hottman et al. (2016); Mayer et al. (2021)) treat multiproduct firms as collections of independent single-product firms. Under this assumption, multiproduct firms can be considered multiple fictitious firms splitting products and analyzed using standard firm-level production network models. We examine the validity of this nonjoint production assumption and determine whether multiproduct firms differ from collections of single-product firms.

To investigate this, We exploit heterogeneous exposure to local buyer shocks for each firm's products and investigate whether firms engage in joint production in the spirit of Ding (2023) but in a production network context. Specifically, We examine how demand shocks to specific products affect the production of other products within the same firm. We use monthly data from 2019 to 2021. We treat different establishments of the same firm as different firms, allowing us to treat regional differences in COVID-19 lockdowns as exogenous shocks. We sum the monthly quantity traded and value for every buyer firm-seller firm-product triplet transaction. We divide the value traded over its quantity to obtain average monthly prices for each triplet.

Our empirical strategy consists of two main components: First, we employ a validation step to confirm that lockdowns indeed caused a decrease in intermediate input purchases, which can be interpreted as demand shocks from the perspective of supplier firms with buyers in lockdown areas. Second, we conduct a main event study examining how these shocks affect the production of other products within the same firm.

COVID-19 Lockdowns in Chile

To provide context for our analysis, Starting in March 2020, the Chilean government declared county-specific lockdowns due to COVID-19. Like many countries, the lockdowns were unexpected; we focus on March 2020, when the shock was unanticipated. Figure 1 shows regions that experienced lockdowns in March 2020, illustrating the sparse distribution of lockdowns due to COVID-19 across the country.

Figure 1: Distribution of Early Covid-19 lockdown in Chile

Notes: Lockdown counties as of March 2020 are red; all others are gray.

Lockdown Effects on Intermediate Input Purchases

Before proceeding to our main analysis, we first establish that COVID-19 lockdowns resulted in reduced intermediate input purchases by buyers in lockdown areas, which can be interpreted as a negative demand shock to suppliers in non-lockdown areas.

We posit that lockdowns reduced intermediate input transactions from firms in unaffected (gray) counties to buyers in counties that experienced early COVID-19 lockdowns (red). We consider this an unanticipated intermediate inputs demand shock to specific products of firms in non-lockdown counties. To test this hypothesis, we estimate the following reduced-form specification at the buyer level:

$$\log M_{it} = \beta \, Lockdown_{it} + FE_t + FE_i + \varepsilon_{it}, \tag{1}$$

where M_{it} denotes total intermediate input purchases of a firm i at time t and $Lockdown_{it}$ is a dummy variable equal to one if firm i's location was under lockdown at time t, and zero otherwise. To address potential bias arising from firms in lockdown areas purchasing from suppliers also in lockdown areas, we report results from a restricted sample including only firms with suppliers in non-lockdown areas.

Table 2: Lockdown and intermediate input purchases

	(1)	(2)	(3)
Lockdown Dummy	-0.222***	-0.230***	-0.191***
	(0.0524)	(0.00521)	(0.0589)
Firm FE	Y	Y	Y
Time FE	N	N	Y
Sector × Time FE	N	Y	N
Restricted sample	N	N	Y
Observations	4,345,534	4,345,534	378,646

Notes: The Table reports the results of estimating equation (1) by ordinary least squares (OLS), clustered at the firm-municipality level. The sample periods are January 2019 to March 2020. Columns (1) and (2) report results for the full sample. Column (3) presents the results restricted to firms with no suppliers in the lockdown area. Three stars indicate statistical significance at the 1% level.

The results confirm our hypothesis: The coefficient of interest, β, is negative, indicating that purchases of intermediate inputs from lockdown counties decreased by about 20% on average. This result confirms that we can interpret the decrease in purchases as a negative demand shock to intermediate inputs sold by firms in non-lockdown regions to buyers in lockdown regions.

Emprical Evidence for Joint Production

Having validated our use of lockdowns as a source of demand shocks, we now focus on the central question of our analysis. We aim to determine whether firms engage in joint production.

To investigate this, we examine how a demand shock to one product affects the production of other products within the same firm. We focus on shocks to a firm's main product, defined as the product with the highest sales from January 2019 to December 2021. We define a firm as experiencing a demand shock to its main product if at least one buyer of its main product is located in an area that implemented a lockdown in March 2020 due to the initial COVID-19 outbreak. This definition focuses on the early lockdowns to ensure the exogenous nature of the shock, as these initial closures were largely unexpected.

To quantify this effect, we study the impact of demand shocks to a firm's main product on the production of its other products using an event-study specification for all products $g \neq m$:

$$\log X_{igt} = \sum_{\substack{j=-11 \\ j\neq -1}}^{10} \beta_j D_{i,t-j} + FE_{ig} + FE_t + \varepsilon_{igt}, \tag{2}$$

where X_{igt} represents either the quantity or price of product g for firm i at time t. $D_{i,t-j}$ is a treatment indicator equal to one if firm i was treated j months ago. FE_{ig} and FE_t are firm-product and time fixed effects, respectively. The coefficients of interest are β_t, which capture the effect of the main product's demand shock on other products' quantities or prices at different time points relative to the shock.

To obtain unbiased estimates of β_j, the treatment indicator $D_{i,t-j}$ must be conditionally orthogonal to the error term ε_{igt}. A key concern is that supply-side shocks could be correlated with the lockdown if suppliers and main product buyers are located in the same area, potentially confounding our results. To address this issue and isolate the impact of demand shocks from the main product while ruling out direct supply shocks, we impose the following restrictions:

1. **Firm Location**: The firm itself is not located in an area under lockdown.

2. **Supplier Location**: The firm's direct suppliers are not subject to lockdown shocks.

3. **Buyer Location for Product** g: None of the buyers of product g are located in lockdown areas.[5]

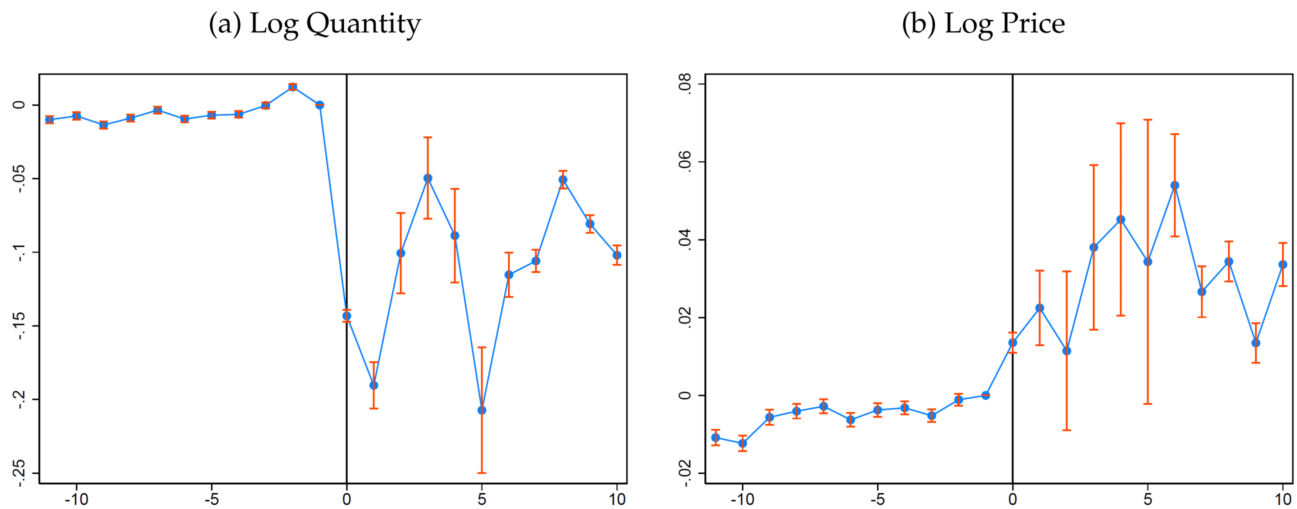
Restrictions 1 and 2 help eliminate direct supply-side effects, ensuring that any observed changes in production are not due to supply disruptions affecting the firm or its suppliers. Condition 3 ensures that product g is not subject to a direct demand shock, allowing us to attribute any changes in its production to the demand shock affecting the main product m.

Our treatment group consists of firms meeting these conditions and experiencing a demand shock to their main product in March 2020. The control group includes firms satisfying the conditions but whose main products did not experience a demand shock

[5] Within the same firm, the set of buyers often differs across products. Therefore, even if the main product's buyers are affected by the lockdown, it does not necessarily imply that buyers of other products are similarly affected. This distinction is further detailed in the Appendix A.1.

until t (i.e., the buyers of the main product have never been in a lockdown area from March 2020 until t). Figure 2 presents the results of regression (2) [6].

Figure 2: The effects of demand shocks to the main product on the production of other products within the firm



Notes: Standard errors are clustered at the firm-county level, and the error bands represent 95% confidence intervals. The X-axis represents the time to treat, with 0 denoting March 2020 when the main product experienced the demand shock, and other values indicating the number of months before or after this event.

First, prior to the shock, the differences in quantities and prices between the treatment and control groups remained stable and close to zero. This observation supports the parallel trends assumption and aligns with the interpretation that the initial closures were unanticipated. Second, we detect a significant quantity effect: at the time of the shock ($t = 0$), there is a 14% decrease in the quantity produced of non-main products when the main product experiences a demand shock. Notably, this effect persists throughout the post-shock period, indicating that the impact is not transitory. Third, we observe an increase in the price of non-main products on impact. This price increase remains elevated above the pre-shock level even ten months after the initial shock. In addition to the quantity response, this price response is consistent with the assumption of joint production, which will be introduced in the next section.

[6] A comparison of observable characteristics between the treatment and control groups is provided in Table A1 of the appendix.

Discussion on Other Within-Firm Spillover Mechanisms

Our empirical results contradict the nonjoint production in multiproduct firms. This outcome suggests that demand shocks to one product affect the production of other products within the same firm. Next, we compare our findings with relevant studies.

First, Almunia et al. (2021) propose a model of diminishing returns to scale or firm-specific factors at the firm level to explain how a decline in domestic demand in Spain affects exports. Their model predicts that when there is a negative demand shock for a product in one market, firm-specific factors are reallocated to another product in another market, positively affecting the production of the same product in other markets. This prediction contrasts with our findings, which show negative spillovers across products within the same firm.

Second, Ding (2023) examines joint production effects in the US using five-year Census data, focusing on industries sharing knowledge-intensive inputs. Their model, like ours, predicts a negative effect. The study shows positive spillovers across industries sharing intangible inputs; however, this mechanism is unlikely to explain our results. The differences in time horizon (five years vs. monthly data) and research and development (R&D) intensity (Chile's R&D spending is less than one-tenth that of the US as a percentage of GDP) limit its applicability to our context.

Another relevant study is that of Giroud and Mueller (2019), who use US multiregion firm data to model demand-driven regional spillovers based on financial constraints. In this model, firms facing credit constraints optimize resource allocation across regions. A negative demand shock in one region reduces employment in other regions due to shared financial constraints within the firm. This mechanism predicts a negative response; a negative shock to one product would lead to reduced production of other products through financial constraints. In addition, although Giroud and Mueller (2019) does not report price reactions, Kim (2020) finds that firms in financial distress reduce product prices by selling off inventory, which is contrary to the results of this paper and suggests a negative price response. In Appendix A.1, we conduct the same event study using a subset of firms unlikely to be under financial constraints. The results do not change significantly, suggesting they are unlikely to be driven solely by financial constraints.

Our results suggest that multiproduct firms are not collections of independent product lines. These empirical results motivate the theoretical framework we develop in the following section.

3 A Theory to Aggregate Distortions in Networks with Multiproduct Firms

We propose a theoretical framework to analyze resource misallocation in multiproduct firms' production networks. This framework generalizes the work of Hulten (1978) and Baqaee and Farhi (2020) and rationalizes our empirical findings on the prevalence and characteristics of multiproduct firms in production networks. We characterize the first-order effects of firms' product-level shocks on aggregate total factor productivity (TFP) growth in an economy with arbitrary wedges, product-level production networks, and joint production.

3.1 Joint Production

Our empirical results show that demand shocks to specific products affect the production of other nonshocked products within the same firm. We adopt a joint production setup, where firms use common inputs to produce different products, simultaneously allowing some inputs to be used in multiple products.

Following Hall (1973), let $J(\boldsymbol{q}, \boldsymbol{x})$ be a joint production function, where $\boldsymbol{q}$ is a vector of outputs and $\boldsymbol{x}$ is a vector of inputs. The joint cost function is derived from the firm's cost minimization problem, as follows:

$$C(\boldsymbol{q}, \boldsymbol{p}) \equiv \min_{x \in V(\boldsymbol{q})} \boldsymbol{p}'\boldsymbol{x},$$

where $V(\boldsymbol{q})$ is the input requirement set, $V(\boldsymbol{q}) = \{x | J(\boldsymbol{q}, \boldsymbol{x}) \geq 0\}$ and $\boldsymbol{p}$ is a vector of input prices. We introduce two assumptions about the shape of a joint production function, which will be used throughout this paper.

Assumption 1. Constant return to scale (CRS): $J(\boldsymbol{q}, \boldsymbol{x}) = 0$ implies $J(\lambda \boldsymbol{q}, \lambda \boldsymbol{x}) = 0$ for any scalar λ.

Unlike a single-output production function, the output is a vector. Note that we do not assume CRS for each single-output production function.

Assumption 2. Separability between Input and Output Bundles: The joint production function can be written as $J(\boldsymbol{q}, \boldsymbol{x}) = -F^Q(\boldsymbol{q}) + F^X(\boldsymbol{x})$, and the joint cost function as $C(\boldsymbol{q}, \boldsymbol{x}) = H(\boldsymbol{q})\varphi(\boldsymbol{p})$.

Note that this differs from assuming nonjoint production functions when a firm is multiproduct, where the output, q, is a single product, not a vector; it degenerates to $F^{Q}(q) = q$. Example 1 illustrates a joint production function satisfying assumptions 1 and 2:

Example 1. Constant Elasticity of Transformation Output Bundle and Constant Elasticity of Substitution Input Bundle (CET-CES):

$$\underbrace{\left(\sum_{g} q_g^{\frac{\sigma+1}{\sigma}}\right)^{\frac{\sigma}{\sigma+1}}}_{\text{Output bundle}} = A \underbrace{\left(L^{\frac{\theta-1}{\theta}} + K^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}}_{\text{Input Bundle}},$$

The associated cost function is

$$C\left(\boldsymbol{q}, w, r\right) = \frac{1}{A}\left(\sum_{g} q_g^{\frac{\sigma+1}{\sigma}}\right)^{\frac{\sigma}{\sigma+1}} \left(w^{1-\theta} + r^{1-\theta}\right)^{\frac{1}{1-\theta}},$$

where L and K are the two inputs; w and r are their prices, and $\boldsymbol{q}$ is a vector of outputs.

The input bundle takes a standard CES function; the output bundle is a vector of products rather than a scalar. The parameter σ is called the constant elasticity of transformation; it gives a constant value to the production possibility frontier's curvature of the products within a firm. This example is illustrative as our theoretical framework requires no parametric assumption.

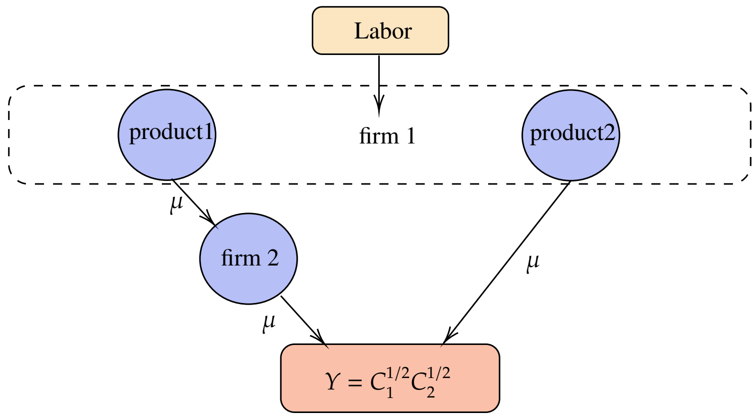
3.2 Parametric Examples of Misallocation in Networks with Joint Production

We provide simplified examples to help gain an intuition about how joint production affects resource allocation and aggregate TFP.

3.2.1 Without joint production

We begin with an example of a production network with multiproduct firms but without joint production. Consider an economy with two firms, as illustrated in Figure 3.

Figure 3: A simplified economy with production networks and multiproduct firms



Firm 1 uses labor to produce two differentiated products using labor (L) as unique input, $q_{11} = L_1$, $q_{12} = L_2$, where $L = L_1 + L_2$. Product 1 is sold to firm 2, while product 2 is sold directly to households. For simplicity, we assume that both products have the same markup, μ. firm 2 uses product 1 from firm 1 as a production input and produces a different product using a linear technology ($q_2 = q_{11}$) that sells to households with markup μ. Final consumption goods are aggregated using a Cobb-Douglas function $Y = c_1^{1/2} c_2^{1/2}$, where $c_1 = q_2$, $c_2 = q_{12}$. In this simple economy, Y is the real GDP, and aggregated TFP can be defined as $TFP = Y/L$.

In a production network environment, distorted resource allocation arises not only from firms' own markups but also from downstream firms' markups. In this example, product 1 is sold to households with double marginalization; firm 1 charges a markup to firm 2, and firm 2 charges a markup to the household. As a result, product 1 from firm 1 suffered from a higher distortion than Product 2 from firm 1, both relative to a competitive setup.

Assume that product 1 from firm 2 receives a shock that reduces its markup ($\Delta \log \mu_{21} = -\epsilon$). This shock increases demand for product 1, which is too small to begin with. By taking a first-order approximation of the change in aggregate TFP, we obtain the following response:

$$\Delta \log TFP = \frac{\mu - 1}{2(1 + \mu)} \epsilon > 0.$$

As product 1's markup decreased, the production of product 1 increased, increasing a portion of the economy that was too small to begin with because of a higher initial markup. Also, product 1's relative price decreased. As a result, the household shifts away from product 2 and increases its consumption of product 1 (indirectly through firm 2 product), which boosts overall economic efficiency (aggregate TFP).

3.2.2 Joint Production

Now, let us introduce joint production into our simplified economy example. Instead of separate production functions for each product, firm 1 now uses a joint production technology to produce both products simultaneously:

$$\left(q_{11}^{\frac{\sigma-1}{\sigma}} + q_{12}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} = L_1.$$

Assume that product 1 from firm 2 receives the same shock as before that reduces its markup ($\Delta \log \mu_{21} = -\epsilon$). By taking a first-order approximation of the change in TFP, we obtain the following response:

$$d \log TFP = \left(\frac{\sigma}{\sigma+1}\right)\left(\frac{\mu-1}{2(1+\mu)}\right)\epsilon > 0.$$

A negative markup shock induces a relative price decline, triggering a substitution effect that shifts household expenditures towards product 1 (indirectly through firm 2 product). Parameter σ governs the elasticity of transformation between the two products of firm 1; how easily firm 1 can adjust its product mix for its two products. The elasticity of transformation will determine the changes in aggregate TFP, which we will illustrate in two extreme cases.

As σ approaches infinity, firm 1 can freely choose the ratio of products in response to a markup decrease to one of its products. The change in aggregate TFP converges to the maximum possible improvement, as firm 1 can freely adjust its production inputs for each of its products.

Conversely, as σ approaches zero, the production technology becomes Leontief in outputs, with fixed proportions. To illustrate this, consider a petroleum refining firm producing gasoline (product 1) and diesel (product 2). Diesel production is constrained by the refining process's specific technological and chemical properties of crude oil; the firm cannot easily switch the oil used to produce gasoline to produce diesel instead.

This rigidity in production means that refineries remain bound to produce gasoline and diesel fuel in fixed proportions despite relative prices or demand changes. Consequently, the potential for production factors reallocation disappears, eliminating aggregate TFP growth changes in response to markup changes.

While an infinite or zero σ are extremes, these examples illustrate that ignoring the rigidities associated with joint production may lead to overestimating aggregate TFP changes in response to shocks, as most literature implicitly assumes σ approaches infinity. However, the degree to which joint production constrains aggregate TFP growth is an empirical question that depends on how easily firms can switch inputs to produce different outputs (the elasticity of transformation in this example), which may vary across firms and industries.

In the next section, we develop a framework to measure these effects in a general production network setting without requiring parametric assumptions about the joint production technology.

3.3 General Network Setup

Regarding input–output notation and definitions, we follow Baqaee and Farhi (2020) to present our generalization and add product-level (instead of firm-level) objects. Without joint production, every product can be considered a fictitious firm so that the Baqaee and Farhi (2020) setup applies.

Multiproduct Firms

Firm $i \in \mathcal{N}$ produces product $g \in \mathcal{G}$ and uses products $g' \in \mathcal{G}$ from other firms $j \in \mathcal{N}$ and factors (Labor, L and Capital, K) as production inputs. [7] We assume the following production set with CRS and separability:

$$F_i^Q\left(\underbrace{\left\{q_{ig}\right\}_{i\in N, g\in G}}_{\text{outputs}}\right) = A_i F_i^X\left(\underbrace{\left\{x_{i,jg'}\right\}_{j\in N, p\in G}}_{\text{Intermediate product } g' \text{ from } j}, L_i, K_i\right), \tag{3}$$

Firms charge a product-specific markup, μ_{ig}, over its product-specific marginal cost; thus, the price is defined as $p_{ig} = mc_{ig}\mu_{ig}$.

[7] We treat factors exhibiting zero return to scale production functions; they generate production inputs without using inputs from other firms.

Final Demand

A representative household with a homothetic utility function of $U\left(c_{ig}, ..., c_{NG}\right)$ receives income from factor payments and profits from firms they own, following a budget constraint:

$$\sum_{i\in\mathcal{N}}\sum_{g\in\mathcal{G}} p_{ig}c_{ig} = \sum_{f\in\{L,K\}} w_f L_f + \sum_{i\in\mathcal{N}}\sum_{g\in\mathcal{G}}\left(1 - 1/\mu_{ig}\right)p_{ig}q_{ig}.$$

Each product can be consumed by final consumers (c_{ig}) or used as an input in production by other firms ($x_{ji,g}$). The following resource constraint applies:

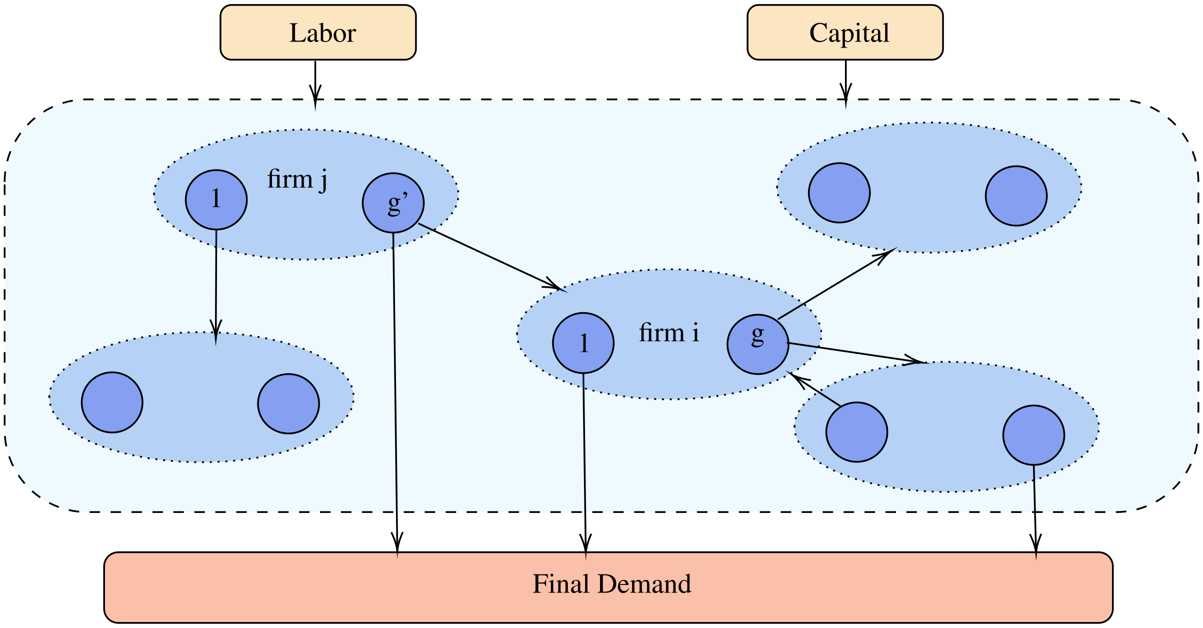
$$q_{ig} = c_{ig} + \sum_{j\in\mathcal{N}} x_{jig}, \quad \sum_{i\in\mathcal{N}} L_i = L, \quad \sum_{i\in\mathcal{N}} K_i = K.$$

Figure 4 presents a stylized representation, showing the flow of products.

General Equilibrium

Given a vector of firm-level productivity, $\boldsymbol{A}$, and vector of product-level markups, $\boldsymbol{\mu}$, for all $i \in \mathcal{N}$ and $g \in \mathcal{G}$, the general equilibrium is a set of prices (p_{ig}) intermediate input choices ($x_{ijg'}$), factor input choices (L_i, K_i), output, (q_{ig}), and consumption choices (c_{ig}). As such, (i) the price of each product is equal to its markup multiplied by its marginal cost, (ii) households maximize utility under budget constraints, given prices, and (iii) markets are clear for all products and factors.

Figure 4: Graphical illustration of networks with multiproduct firms



Notes: The dashed line represents firms' universe $\mathcal{N}$, the dotted circled line represents each firm's boundary, and the circled line represents each product within a firm. The two top nodes represent factors, and the bottom node represents households. Arrows represent the direction of input flows.

3.4 Input–Output Definitions

We introduce notation for input–output objects to state our decomposition results.

Product-Level Input–Output Matrix

The product-level input–output matrix $\tilde{\Omega}$ is a $(\mathcal{N}\mathcal{G} + \mathcal{F})$ square matrix. Here, $\mathcal{N}$ is the number of firms, $\mathcal{G}$is the number of products, and $\mathcal{F}$ is the number of factors. $\tilde{\Omega}$ has at its ig, jg'^{th} element the expenditure share of product g' from firm j and factor $f \in \mathcal{F}$ used by firm i in production over firm i total costs (of producing all its products). The separability assumption indicates that the same expenditure share applies for all products, g, that firm i produces; thus, $\tilde{\Omega}_{ig,jg'}$ and $\tilde{\Omega}_{ig,f}$ are as follows.

$$\tilde{\Omega}_{ig,jg'} = \frac{p_{jg'}x_{i,jg'}}{\sum_{j,p} p_{jg'}x_{i,jg'} + \sum_f w_f L_{if}}, \quad \tilde{\Omega}_{ig,f} = \frac{w_f L_{if}}{\sum_{j,p} p_{jg'}x_{i,jg'} + \sum_f w_f L_{if}}.$$

The product cost-based Leontief inverse $\tilde{\Psi}$ captures each firm-product pair's direct and indirect cost exposures through production networks. We use each $\tilde{\Psi}$ element to measure the weighted sum of all paths between two nonzero firm-product pairs.

$$\tilde{\Psi} \equiv (I - \tilde{\Omega})^{-1} = i + \tilde{\Omega} + \tilde{\Omega}^2 + \dots$$

We define the value-added share vector, b, as follows:

$$b_{ig} = \begin{cases} \frac{p_{ig} c_{ig}}{GDP} & \text{if } i \in \mathcal{N}, g \in \mathcal{G} \\ 0 & \text{otherwise} \end{cases}$$

We set GDP to be the numeraire and define the product-level cost-based Domar weight, $\tilde{\lambda}_{ig}$. [8] This measures the importance of product g from firm i in final demand in two dimensions: directly when sold to final consumers and indirectly through the production network when product g is sold to other firms and eventually reaches final consumers via downstream production networks.

$$\tilde{\boldsymbol{\lambda}}' \equiv \boldsymbol{b}'\tilde{\Psi} = \boldsymbol{b}' + \boldsymbol{b}'\tilde{\Omega} + \boldsymbol{b}'\tilde{\Omega}^2 + \dots$$

Firm-Level Aggregation

Summing over products by firms allows us to recover the firm-level cost-based Domar weight $\tilde{\lambda}_i$, which we use to compute the within-firm product-level Domar weight share s_{ig}:

$$\tilde{\lambda}_i = \sum_{g \in \mathcal{G}} \tilde{\lambda}_{ig}, \quad s_{ig} = \frac{\tilde{\lambda}_{ig}}{\tilde{\lambda}_i}.$$

Finally, we define firm-level aggregate markup as follows:

$$\mu_i = \frac{\text{sales of } i}{\text{total cost of i}},$$

National Accounts

GDP is defined as the sum of all product values consumed by final consumers: $GDP = \sum_{i \in \mathcal{N}} \sum_{g \in \mathcal{G}} p_{ig} c_{ig}$. Real GDP (Y) changes can be computed as follows:

$$d \log Y = d \log GDP - \sum_{i \in \mathcal{N}} \sum_{g \in \mathcal{G}} \frac{p_{ig} c_{ig}}{GDP} d \log p_{ig}.$$

[8] We denote $\tilde{\Lambda}_f$ with $f \in \{L, K\}$.

Factor shares are defined as

$$\Lambda_L = \frac{wL}{GDP}, \quad \Lambda_K = \frac{rK}{GDP}.$$

3.5 Cumulative Markup

We now define the cumulative markup, Γ_{ig}, which is the critical input to our proposed sufficient statistic strategy. The cumulative markup is defined as the ratio of the product cost-based Domar weight, $\tilde{\lambda}_{ig}$, to the product sales share (to GDP), adjusted by the product-level markup.

Definition 1. Cumulative Markup

$$\Gamma_{ig} \equiv \underbrace{\frac{\tilde{\lambda}_{ig}}{\textit{sales share}_{ig}}}_{\text{downstream distortion}} \times \underbrace{\mu_{ig}}_{\text{own markup}},$$

The cumulative markup summarizes the cumulative distortion in the downstream supply chain of product g sold by firm i. In efficient economies, with no markups, the product cost-based Domar weight equals observed sales shares, generating a cumulative markup equal to one for all products and firms. Conversely, in an inefficient economy, a portion of the indirect demand transmitted from downstream firm-product pairs to upstream firm-product firms is absorbed as profit by downstream firms. This effect accumulates in each supply chain transaction upstream until indirect demand reaches product g sold by firm i; thus, the sales share of a product is smaller relative to an efficient economic outcome. Therefore, the larger the ratio, the greater the cumulative distortions in the downstream supply chain.

Next, we define the relative product cumulative markup, which ranks product distortions within a given firm. It is measured as the relative downstream distortion of product g with respect to the average distortion of all products within firm i, Γ_i.

Definition 2. Relative Cumulative Markup

$$\gamma_{ig} \equiv \frac{\Gamma_{ig}}{\Gamma_i},$$

where the average distortion of firm i is defined as: $\Gamma_i \equiv \sum_g \tilde{\lambda}_{ig} / \sum_g \left(\textit{salesshare}_{ig}/\mu_{ig}\right)$.

A Simple Example of Cumulative Markup

Let's revisit our simplified economy composed of two firms and a representative household to illustrate the concept of cumulative markups and their computation. We normalize GDP to 1; hence, in this setup, sales (shares) to final consumption are (1/2) for products 1 and 2; however, firm 1's sales of product 1 are reduced by the markup charged by firm 2, which is $(1/2)/\mu$.

The product cost-based Domar weights are 1/2 for products 1 and 2. In matrix notation, the value-added share vector (b) and the product cost-based input-output matrix ($\tilde{\Omega}$) are:

$$b = \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{bmatrix}, \quad \tilde{\Omega} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where the matrix and vector components are arranged in the following order: product 1 and 2 of firm 1, firm 2, and labor. Therefore, the product cost-based Domar weights can be computed as:

$$\begin{aligned} \tilde{\boldsymbol{\lambda}}' &= \boldsymbol{b}' + \boldsymbol{b}'\tilde{\Omega} + \boldsymbol{b}'\tilde{\Omega}^2 + \dots \\ &= \underbrace{\begin{bmatrix} 1/2, & 0, & 1/2, & 0 \end{bmatrix}}_{\text{Final demand}} + \underbrace{\begin{bmatrix} 0, & 1/2, & 0, & 0 \end{bmatrix}}_{\text{Indirect demand}} \\ &= \begin{bmatrix} 1/2, & 1/2, & 1/2, & 0 \end{bmatrix} \end{aligned}$$

These weights represent the counterfactual sales shares if markups were removed while keeping expenditure shares constant. Following the definition, the cumulative markup is $\Gamma_{ig} = \frac{\tilde{\lambda}_{ig}}{\text{sales share}_{ig}} \times \mu_{ig}$ for firm 1:

$$\Gamma_{11} = \frac{1/2}{(1/2)/\mu}\mu = \mu^2, \qquad \Gamma_{12} = \frac{1/2}{(1/2)}\mu = \mu.$$

The following Table 3 summarizes the results.

Table 3: Sales share, cost-based Domar weight, and cumulative markup in this example

	product 1 of firm 1	product 2 of firm 1
(1) Sales share	$(1/2)/\mu$	$1/2$
(2) Cost-based Domar weights	$1/2$	$1/2$
(3) Cumulative Markup: $(2)/(1) \times$ own markup	μ^2	μ

The markup of product 2 from firm 1 and the product from firm 2 equal μ. Comparatively, product 1 from firm 1 has a larger cumulative markup of μ^2 than that of product 2, reflecting the product's markup and the downstream distortions the product faces. In this case, product 1 from firm 1 generates a distortion by charging a markup; it is subject to an additional distortion downstream production networks as firm 2 uses a marked-up input on its production. The sum of both distortions is the main driver when characterizing the multiproduct channel of allocative efficiency, which the next section discusses.

3.6 Aggregation Theorem with Multiproduct Firms within Production Networks

This section generalizes the concept of an allocation matrix introduced by Baqaee and Farhi (2020) to allow for joint production and derive sufficient statistics.

Let $\mathcal{X}$ be an $(\mathcal{N}+\mathcal{F}) \times (\mathcal{N}\mathcal{G}+\mathcal{F})$ admissible input allocation matrix; the columns are buyer firms, and the rows are seller-product pairs. Each element, $\mathcal{X}_{ijg} = \frac{x_{ijg}}{q_{jg}}$, is the share of the output of product g produced by firm j that firm i uses as a production input.

A productivity shock ($\mathrm{d}\log A$) and a markup shock ($\mathrm{d}\log \mu$) effect in real GDP, Y, can be decomposed into a pure change in productivity ($\mathrm{d}\log A$) for a given fixed allocation matrix $\mathcal{X}$ and changes in the distribution of $\mathcal{X}$ ($\mathrm{d}\mathcal{X}$) holding productivity constant. In vector notation:

$$\mathrm{d}\log \mathcal{Y} = \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \log A}\mathrm{d}\log A}_{\Delta \text{ Technology}} + \underbrace{\frac{\partial \log \mathcal{Y}}{\partial \mathcal{X}}\mathrm{d}\log \mathcal{X}}_{\Delta \text{ Allocative Efficiency}}. \tag{4}$$

We now present a decomposition of changes in aggregate TFP that considers multiproduct firms and arbitrary production networks with product-level distortions.

Proposition 1. *Growth Accounting in Production Networks with Multiproduct Firms: To a first* *order, aggregate TFP can be decomposed into technology and allocative efficiency terms as follows:*

$$d\log TFP = \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log A_i}_{\Delta\ Technology} + \underbrace{\underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i Cov_{\boldsymbol{s}_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i}\right)}_{Multiproduct\ term} - \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log \mu_i}_{Firm\text{-}level\ Markup} - \underbrace{\sum_{f\in\mathcal{F}} \tilde{\Lambda}_f d\log \Lambda_f}_{Aggrregate\ Factor\ Shares}}_{\Delta\ Allocative\ Efficiency}, \quad (5)$$

where $\boldsymbol{\gamma}_i = (\gamma_{i1}, ..., \gamma_{iG})$ *and* $d\log \boldsymbol{p}_i = (d\log p_{i1}, ..., d\log p_{iG})$.

Appendix D presents the proof. The change in aggregate TFP can be decomposed into technology and allocative efficiency terms. The technology term represents a weighted average of changes in firm-level Hicks-neutral productivity using cost-based Domar weights. The allocative efficiency term is further decomposed into a multiproduct firm term, a change in aggregate factor shares, and firm-level average markup changes.

The multiproduct term captures the allocative efficiency implications of firm-level product mix adjustments. The contribution to each product's allocative efficiency depends on the product's price relative to the firm's average price change and the product's relative cumulative markup in the product's product mix. In economies with joint production, the opportunity cost of changing the product mix is generally not constant. This means that it may vary from firm to firm and product pair to product pair, affecting the product level unit cost depending on the curvature of the production possibility frontier. For example, an oil refinery producing gasoline and diesel may face different tradeoffs when adjusting its production mix compared to a dairy farmer producing milk and meat. The observed relative price changes per product contain all the information on the production possibility frontier for this firm. These price changes reflect the firm's internal tradeoffs in production. The implication for allocative efficiency of the price change associated with a change in the product mix can then be computed using the covariance of the relative price changes and the cumulative markups of the product whose price has changed. For example, suppose the price change for a given product g in firm i is higher than the average price change for firm i, and product g is more distorted than the average product within firm i. In this case, the covariance will be positive. Intuitively, a decrease in the price of an upstream product absorbs the distortion for all downstream firms that use that product as an input (directly or indirectly). This is because lower input prices reduce the impact of markups along the supply chain. The covariance implies that this effect increases when prices for relatively more distorted products fall. To calculate the aggregate effect across the economy, we can sum these firm-level covariances using Domar weights, which indicate its macroeconomic importance. This aggregation allows us to quantify the overall impact of product mix changes on allocative efficiency in the economy.

Regarding the factor shares and firm-level markup terms, if the initial equilibrium is inefficient, the products charging markups are underproduced relative to an efficient economy. Improving the allocation involves reallocating resources to a more distorted part of the economy, such as firms' product pairs that charge relatively high markups. A decrease in factor shares implies reallocating resources to the portion of the economy with relatively high markups; however, if the change in factor share is due to a change in markup, this is a mechanical change and does not imply reallocation. Therefore, the contribution of the change must be purged, which the firm-level markup term captures. The factor shares and firm-level markup terms are proposed by Baqaee and Farhi (2020). Both terms are valid under a joint production approach and, together with the multiproduct term this work introduces, constitute allocative efficiency.

Relation to Existing Aggregation Theorems

Proposition 1 nests existing aggregation theorems for production networks as a special case.

Corollary 1. *Baqaee and Farhi (2020): If no firms engage in joint production and impose the same markup on all their products (single-product firms assumption), then to a first order, aggregate TFP growth can be decomposed into technology and allocative efficiency terms as follows.*

$$d\log TFP = \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log A_i}_{Technology} \underbrace{- \sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log \mu_i - \sum_{f\in\mathcal{F}} \tilde{\Lambda}_f d\log \Lambda_f}_{Allocative\ Efficeincy}.$$

The proof follows from the fact that the covariance term from proposition 1 is zero because the change in marginal cost and markup for all products within a firm are equal.

Our approach quantifies misallocation through the multiproduct channel by measuring deviations from the single-product, single-markup assumption when product-level data are available; if this assumption holds, the multiproduct term becomes zero. The assumption of uniform marginal costs and markups is unlikely to hold in practice; however, its quantitative relevance remains an empirical question. Our decomposition quantifies the extent to which this assumption is violated and isolates the impact of existing misallocation literature.

Finally, without markups, when prices equal marginal costs, allocative efficiency converges to zero. In this case, all aggregate TFP changes are attributed to technology, aligning with Hulten (1978).

Corollary 2. *Hulten (1978): Growth Accounting in an Efficient Economy:*

$$d\log TFP = \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log A_i}_{\text{Technology}}.$$

The proof follows from the fact that the markup is always 1, the markup change term is 0, and the sum of factor shares is always 1. Therefore, the sum of factor changes is always 0, and the covariance of the multiproduct term is 0 because $\boldsymbol{\gamma}_i = (\gamma_{i1}, ..., \gamma_{iG})$ are all 1 in an efficient economy.

Proposition 1's formula converges to Hulten's theorem when the economy is efficient. Measured aggregate TFP growth equals the Domar weighted sum of firm-level productivity changes.

Applying Our Formula to the Simple Example

Let's revisit our simple example to illustrate how we can apply our decomposition formula without knowing the specific joint production structure. Appendix D provides the proofs for this example. Consider the economy where firm 1 employs workers to produce two differentiated products using a joint production technology:

$$\left(q_{11}^{\frac{\sigma+1}{\sigma}} + q_{12}^{\frac{\sigma+1}{\sigma}}\right)^{\frac{\sigma}{\sigma+1}} = L_1.$$

Applying our decomposition, we find the following steps. First, we compute the multiproduct term as follows:

$$Cov_{\boldsymbol{s}_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i}\right) = -\left(\frac{1}{\sigma+1}\right)\left(\frac{\mu-1}{2(1+\mu)}\right)\epsilon > 0.$$

The single product term is:

$$-\underbrace{\sum_{i\in\mathcal{N}}\tilde{\lambda}_i d\log\mu_i}_{\text{Firm-level Markup}}-\underbrace{d\log\Lambda_f}_{\text{Aggregate Labor Shares}}=\left(\frac{\mu-1}{2\left(1+\mu\right)}\right)\epsilon.$$

Therefore,

$$\Delta\log TFP=\frac{\sigma}{\sigma+1}\left(\frac{\mu-1}{2\left(1+\mu\right)}\right)\epsilon>0.$$

We recover the results from subsection 3.2.2 using Proposition 1. Importantly, we can see that the multiproduct term partially offsets the single-product term. The single-product term suggests a positive reallocation effect due to the markup reduction. However, the multiproduct term, which is negative, dampens this effect. This dampening occurs because joint production constrains the firm's ability to freely adjust its product mix in response to the markup shock. The net effect on TFP is still positive but smaller than what would be predicted by considering only the single-product term.

This example illustrates that ignoring the rigidities associated with joint production may lead to overestimating the TFP contribution of misallocation. Our sufficient statistics automatically detect whether each firm engages in joint production without the need to identify a parametric elasticity of transformation. Some firms may operate independent product lines, while others face strict constraints in adjusting their product mix. The relative price variations of products sold to different buyers within the firm contain all relevant information about the cost of exchanging two goods.

4 Data and estimation

Our analysis relies on a dataset that covers the universe of formal firms operating in Chile from 2016 to 2022; the data are from the Chilean Internal Revenue Service (Servicio de Impuestos Internos, SII). The Chilean tax system requires all formal firms to declare their invoices for transactions with other firms, which provides detailed information on every product, quantity, and price traded. Additionally, we use tax accounting declarations, which provide monthly data on each firm's revenue and input expenditures, including capital and labor costs. A key advantage of the SII data is the use of unique identifiers for firms and workers, enabling the merging of individual and firm data across datasets. We utilize four distinct sources from SII.

The first source used is the value-added tax form, including gross monthly firm sales, materials expenditures, and investment.

Second, the SII provides information from a matched employer–employee census of Chilean firms from 2005. Specifically, firms must report all payments to individual workers, including the sum of taxable wages, overtime, bonuses, and any other labor earnings for each fiscal year. All legal firms must report to the SII; thus, the data cover the total labor force with a formal wage contract, representing roughly 65% of employment in Chile. For any given month, it is possible to identify an individual worker's employment status, their average monthly labor income that year, a monthly measure of total employment, and the distribution of average monthly earnings within the firm.

Third, income tax form data includes yearly information on all sources of a firm's income and expenses. This form allows for computing every individual's actual tax payments for each year. Details on sales and employment are available on this form; however, we use only data on capital stock for each firm and year. This approach allows us to build perpetual inventories using data from the monthly F22 form. We obtain the user cost of capital by multiplying nominal capital stock by the real rental rate of capital, which is built using publicly available data. We use the 10-year government bond interest rate minus expected inflation plus the external financing premium. Finally, we use the capital depreciation rate from the LA-Klems database.

The fourth source comprises electronic tax documents from 2016 onward. These documents provide information on each product (price and quantity) traded domestically or internationally with at least one Chilean firm. We only use domestic transactions and observe the firm-to-firm transactions and the firm's total sales (including firm-to-firm and firm-to-consumer sales). We compute firm-specific product shares for firm-to-firm transactions and assume that their distributions are equivalent to firm-to-consumer transactions to recover the complete distribution of firm sales by product. Each firm-to-firm transaction includes a "detail" column that records the name of each product.

Building on the data cleaning process described in section 2, we process the data to construct product code-level output and input-price indices for each firm using standard Tornqvist indices. We aggregate products into a 290 product code identifier to facilitate comparison between firms, allowing us to estimate product production functions that use the same product across firms.

4.1 Data Cleaning and Implementation Strategy

We begin the data processing by applying filters to the raw data to obtain the final database for empirical analysis. We define a firm as a taxpayer with a tax ID, positive sales, positive materials, positive wage bill, and capital for any given year. We exclude firms that hire less than two employees a year or have capital valued below US$20 in a year. All variables are winsorized at the 1% and 99% levels to mitigate measurement error.

We selected 2016 as the base year for price indices because it was the first year we observed prices for firm-to-firm transactions. This method is widely recognized for estimating aggregate production functions at the firm or plant level when price data is accessible. We use crosswalks developed at the Central Bank of Chile (Acevedo et al. (2023)) to address the challenge of product aggregation (from around 15 million products to 290 product codes). We create aggregated product-level quantity produced and material usage indices, matching product descriptions and characteristics to ensure consistency across firms and over time. Table A2 presents the sample's aggregate statistics, showing the product-level sales ratio to cost-based Domar weights and markups. The first two can be observed directly from the data without imputation; however, the distortion of the product itself, i.e., the markup, must be estimated. For markup estimation, we follow Dhyne et al. (2022), who developed an estimation method for multiproduct production functions assuming joint production. This approach captures the inefficiencies of production factors allocation within a firm, revealing how production changes of one product affect the production of other products within a firm.[9]

4.2 Construction of Sufficient Statistics

We measure five distinct objects to implement the growth accounting framework that includes the multiproduct channel: (1) product-level cost-based Domar weights $\tilde{\lambda}$, (2) product-firm level price indices, (3) product-level markups μ, (4) cumulative markup, and (5) aggregate objects. We discuss each of these in the following subsection.

4.2.1 Product-Level Cost-Based Domar Weights

The product cost-based Domar weights can be calculated using the following equation:

[9] We keep a record of product-specific prices and quantities to build price indices of the composite product.

$$\tilde{\lambda}' \equiv \boldsymbol{b}'\tilde{\Psi} = \boldsymbol{b}' + \boldsymbol{b}'\tilde{\Omega} + \boldsymbol{b}'\tilde{\Omega}^2 + \dots$$

To compute these weights, we must measure value-added shares ($\boldsymbol{b}$) and the input–output matrix ($\tilde{\Omega}$). We measure these objects directly from the data.

Final expenditure shares ($\boldsymbol{b}$) are represented by a vector of dimension $(\mathcal{N}\mathcal{G} + \mathcal{F}) \times 1$. Here, $\mathcal{N}$ is the number of firms, $\mathcal{G}$ is the number of products, and $\mathcal{F}$ is the number of factors. The first $\mathcal{N}\mathcal{G}$ entries are calculated as the residual between a firm product's total sales and its intermediate sales to other firms (measured from the firm-to-firm data). This approach provides a theory-consistent measure of final expenditures. The final $\mathcal{F}$ entries are set to zero, as households do not directly purchase factors. Using firm-to-firm records and factor expenditures, we construct the input–output matrix $\tilde{\Omega}$ at the product-firm level.

Specifically, we compute an annual cost-based input–output matrix by product. We calculate the denominator of each element (indexed by ig, jg') by summing a firm's purchases from all its suppliers, its wage bill, and its capital multiplied by the relevant user cost rental rate of capital. The last two elements of the matrix have wage bills and capital expenditures as their numerators.

The resulting $\tilde{\Omega}$ is a $(\mathcal{N}\mathcal{G} + 2) \times (\mathcal{N}\mathcal{G} + 2)$ matrix that can be expressed as follows:

$$\tilde{\Omega} = \left[\begin{array}{ccc|cc} \tilde{\Omega}_{11,11} & \cdots & \tilde{\Omega}_{11,\mathcal{N}\mathcal{G}} & \tilde{\Omega}_{11,\mathcal{N}\mathcal{G}+1} & \tilde{\Omega}_{11,\mathcal{N}\mathcal{G}+2} \\ & \ddots & & & \\ \tilde{\Omega}_{\mathcal{N}\mathcal{G},11} & & \tilde{\Omega}_{\mathcal{N}\mathcal{G},\mathcal{N}\mathcal{G}} & \tilde{\Omega}_{\mathcal{N}\mathcal{G},\mathcal{N}\mathcal{G}+1} & \tilde{\Omega}_{\mathcal{N}\mathcal{G},\mathcal{N}\mathcal{G}+2} \\ \hline 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 \end{array}\right]$$

Based on the separability assumption, the same expenditure share applies to all products g that firm i produces. The expressions for $\tilde{\Omega}_{ig,jg'}$ and $\tilde{\Omega}_{ig,f}$ are as follows:

$$\tilde{\Omega}_{ig,jg'} = \frac{p_{jg'}x_{i,jg'}}{\sum_{j,p} p_{jg'}x_{i,jg'} + \sum_f w_f L_{if}}, \quad \tilde{\Omega}_{ig,f} = \frac{w_f L_{if}}{\sum_{j,p} p_{jg'}x_{i,jg'} + \sum_f w_f L_{if}}.$$

Factors do not require inputs; thus, the last row of the matrix is zero.

After calculating the product-level cost-based Domar weights, we sum them for the same firms to compute the firm-level cost-based Domar weights and their shares. These will be inputs for Proposition 1.

$$\tilde{\lambda}_i = \sum_{g \in \mathcal{G}} \tilde{\lambda}_{ig}, \quad s_{ig} = \frac{\tilde{\lambda}_{ig}}{\tilde{\lambda}_i}.$$

4.2.2 Product-Firm Level Price Indices

We observe prices for each transaction and aggregate them into 290 product categories. We construct two types of price indices: output and input price indices. We compute firm-product-specific annual price indices for the output price index, which is an input to sufficient statistics and deflates product output for production function estimation. The original data are at the "detail" product level, which we aggregate to a Tornqvist index for each 290 product category the firm owns. Specifically, we construct the following price index:

$$\Delta \log P_{igt} = \sum_{d \in g} \frac{s_{idt} + s_{idt-1}}{2} \Delta \log P_{idt},$$

where d is the detailed category belonging to the upper product category (290 product codes). $\Delta \log P_{idt}$ is the price change, and s_{idt} is the share at time t in the continuing products in category g. We construct our price index with 2016, the starting year of the data, as the base year. We also construct an input price index to deflate material costs for production function estimation. We define one aggregate index per firm since aggregate materials are used as inputs in production function estimation. The construction method is the same as for the output price index.

4.2.3 Product-Level Markups

We estimate product-level markups using the production approach based on Dhyne et al. (2022). This method extends the Ackerberg et al. (2015) production function estimation technique to a multiproduct setting. This approach considers joint production, where firms simultaneously use common inputs to produce multiple products. The approach relies on cost minimization principles to identify unobserved marginal costs for each firm's product. We employ a Cobb–Douglas production function with three factors: capital, labor, and materials. Our results show a product-level markup median of 1.22. Please refer to Appendix C for a detailed explanation of the methodology.

4.2.4 Cumulative Markup

We require product cost-based Domar weights, product sales shares, and product markups to construct the cumulative markup measure:

$$\Gamma_{ig} \equiv \underbrace{\frac{\tilde{\lambda}_{ig}}{salesshare_{ig}}}_{\text{Downstream distortion}} \times \underbrace{\mu_{ig}}_{\text{own markup}} .$$

While the first two components are directly observable in our data, the markup requires estimation. As discussed in Section 3, the ratio of cost-based Domar weights to sales share represents the cumulative distortion accumulated downstream of a product. The cumulative markup, Γ, is essential in constructing the multiproduct term, and it is vital to understand whether this variation arises from downstream distortions or a product's markup.

Table 4: Variance decomposition of $\log \Gamma$

Year	Downstream distortions	Own markup	Covariance
2016	103.3%	0.6%	-3.9%
2017	102.3%	0.7%	-3.0%
2018	102.5%	0.6%	-3.1%
2019	102.8%	0.6%	-3.5%
2020	103.2%	0.7%	-3.8%
2021	103.7%	0.6%	-4.3%
2022	104.9%	0.7%	-5.6%

Notes: We compute the variance decomposition of the logarithm of $\Gamma_{ig} \equiv \frac{\tilde{\lambda}_{ig}}{salesshare_{ig}} \times \mu_{ig}$ for each year. $Var\left(\log \Gamma_{ig}\right) = Var\left(\log\left(\tilde{\lambda}_{ig}/salesshare_{ig}\right)\right) + Var\left(\log \mu_{ig}\right) + 2Cov\left(\log\left(\tilde{\lambda}_{ig}/salesshare_{ig}\right), \log \mu_{ig}\right)$. The first term on the right-hand side is the variance of downstream distortions. The second term is the variance of their own markup, and the last is the covariance of both. We report the percentage of each term on the right-hand side that explains the total variance.

Table 4 presents the variance decomposition of Γ by year. The results show that most variation in Γ stems from downstream distortions, with minimal contribution from the product's markup. This finding is unsurprising, given that downstream distortions represent cumulative wedges throughout the downstream supply chain of the entire economy. In contrast, μ represents a product's own markup. This result implies that the downstream distortions faced by each pair of firms and products are highly heterogeneous when considering product- and firm-level production networks.

When we apply Proposition 1 to the data, the cumulative markup (the multiproduct term using Γ as input) will be less sensitive to mark up estimates. The following section presents results using Γ without the markups and demonstrates our findings' robustness when we approximate the model around a state with an initial common markup within the firm.

Ranking of Cumulative Markups

The analysis reveals that cumulative markups primarily represent accumulated downstream distortions rather than product-level markups. We ranked products by their cumulative markup to better understand which products face increased downstream distortion. Below, we describe and discuss the major product categories. Appendix B presents the complete list of the 30 top and bottom items.

The product categories with the greatest (downstream) distortions mainly comprise business services. For example, insurance brokerage services top the list, followed by employment services (recruitment and supply), electricity distribution to businesses, and postal and courier services. These products are usually upstream inputs that other firms use in production, suggesting insufficient size as distortions accumulate through the supply chain before they reach final demand. An important exception is tobacco, a product close to final demand but ranked high (15th) because of the large wedge accumulated (59.7% tax rate vs 19% VAT tax for other products).

Conversely, the least distorted products include cakes, beer, pet food, personal services such as hospitals, and minerals (copper, silver, and molybdenum), Chile's primary export industry. These products are common downstream products close to Chile's final demand. As a result, the number of supply chains that reach the final consumer is relatively small, and inefficiencies are relatively less likely to accumulate.

4.2.5 Aggregate Objects

In addition to product cost-based Domar weights, markups, and product distortions, we must measure aggregate objects to implement the sufficient statistics presented in Proposition 1. In particular, Y, L, K, Λ_L, and Λ_K denote aggregate value added, employment, capital, and factor shares, respectively. We measure Y, L, and K as the sum of value added, employment, and capital, respectively, for all firms in the economy. Factor shares of GDP, Λ_L and Λ_K, are measured as total compensation and capital with user cost of capital divided by GDP. Real GDP is calculated by deflating GDP with the official GDP deflator.

5 Application: Decomposing Aggregate TFP Growth

This section applies Proposition 1 to analyze aggregate TFP growth for the Chilean economy. Our analysis covers 2016 to 2022, during which Chile's aggregate TFP stagnated and decreased at the margin. This productivity trend aligns with the pattern of productivity stagnation observed in Chile with different computation methods. [10]

We begin by presenting results using the standard assumption in the literature of single-product firms. If firms produces a single product, then Corollary 1 applies:

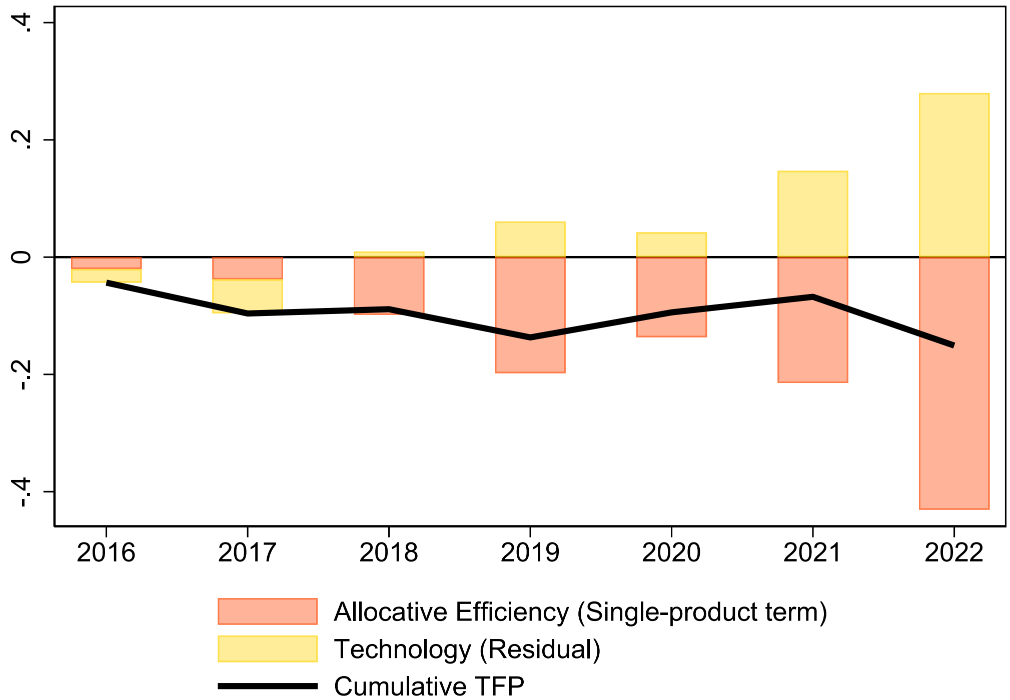
$$d \log TFP = \underbrace{\sum_{i \in \mathcal{N}} \tilde{\lambda}_i d \log A_i}_{\text{Technology}} \underbrace{- \sum_{i \in \mathcal{N}} \tilde{\lambda}_i d \log \mu_i - \sum_{f \in \mathcal{F}} \tilde{\Lambda}_f d \log \Lambda_f}_{\text{Allocative Efficeincy}} .$$

This approach implements growth accounting while overlooking multiproduct firms engaging in joint production. Figure 5 illustrates the decomposition of cumulative changes in aggregate TFP from 2016 to 2022 under this assumption.

Figure 5 shows that the allocative efficiency term (in red) declined over this period. This outcome suggests that high-markup firms contracted further, resulting in a negative reallocation effect; however, the contribution of allocative efficiency exceeds that of the technology (residual) component, particularly during the COVID-19 pandemic and the subsequent high inflation period. To rationalize this disparity, the technology term, measured as a residual, must have increased by around 20%.

[10] CNEP (2023)

Figure 5: Cumulative TFP growth decomposition: Ignoring multiproduct term



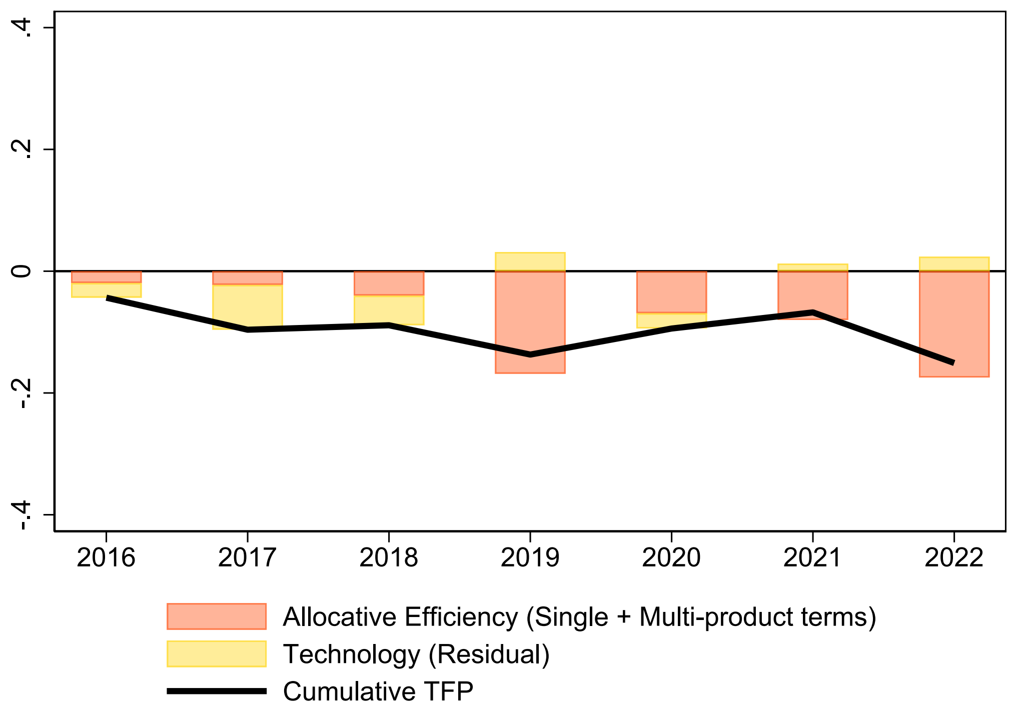
Notes: This Figure shows the cumulative change calculated by applying Corollary 1 repeatedly each year. Technology (residual) is calculated by subtracting allocative efficiency from TFP growth.

Next, we incorporate the multiproduct term using Proposition 1:

$$d\log TFP = \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log A_i}_{\Delta\text{ Technology}} + \underbrace{\underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i Cov_{s_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i}\right)}_{\text{Multiproduct term}} - \underbrace{\sum_{i\in\mathcal{N}} \tilde{\lambda}_i d\log \mu_i}_{\text{Firm-level Markup}} - \underbrace{\sum_{f\in\mathcal{F}} \tilde{\Lambda}_f d\log \Lambda_f}_{\text{Aggrregate Factor Shares}}}_{\Delta\text{ Allocative Efficiency}} . \tag{6}$$

Figure 6 presents the results incorporating the multiproduct term, which reduces the magnitude of the technology (residual) observed in Figure 5. In other words, the multi-product and single-product misallocation terms account for a larger portion of aggregate TFP movements during the COVID-19 pandemic and the resultant high inflation period.

Figure 6: Cumulative TFP growth decomposition with multiproduct term

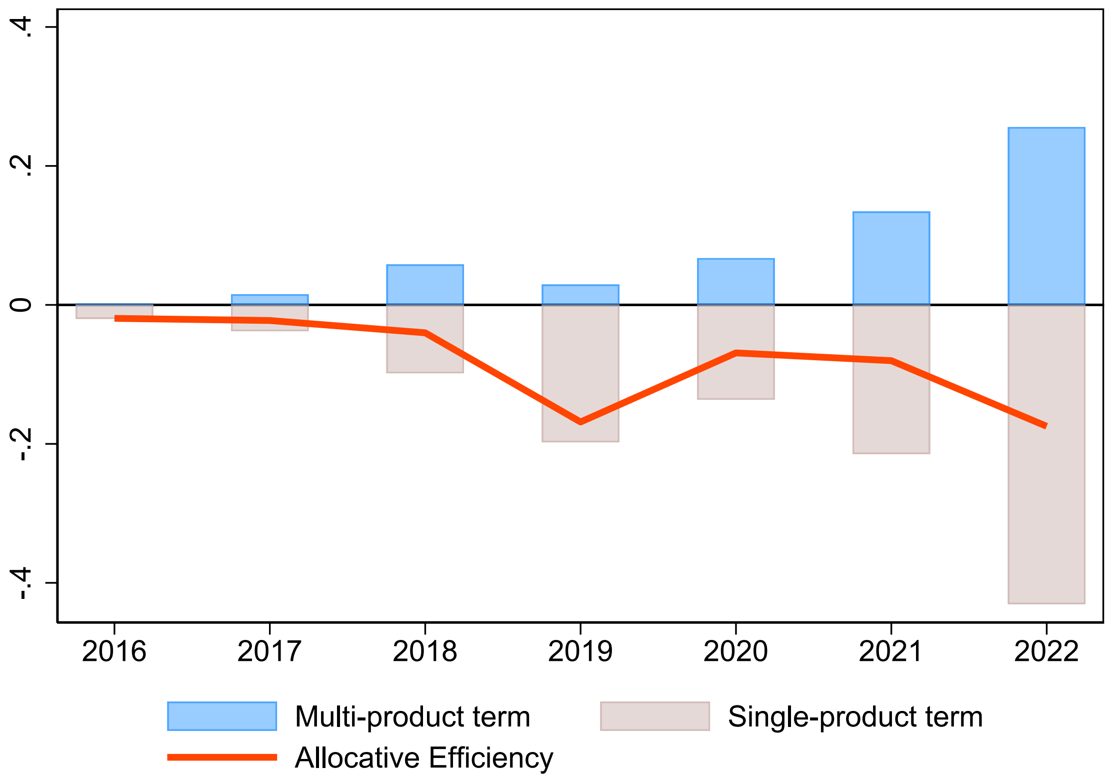


Notes: This Figure shows the cumulative log change calculated by repeatedly applying the equation from Proposition 1 each year. Technology (residual) is calculated by subtracting allocative efficiency from aggregate TFP growth.

Reallocation effects considering joint production explain 86% of the observed aggregate TFP growth. Conversely, as shown in Figure 5, ignoring joint production leads to overestimating resource misallocation. This result suggests that considering joint production considerably decreases the reallocation implied under the traditional assumption that firms produce only single products.

This finding is consistent with the joint production mechanism described in Section 3. When firms engage in joint production, they create multiple products using common inputs. When a given product receives a shock, if firms face technological constraints to adjust their product mix (non-infinite elasticity of transformation), firms will struggle to reallocate productive resources from one product to another. The reallocation through substitution among products within multi-product firms is dampened, and reallocation is not materialized to the extent suggested under the single-product firm assumption.

Figure 7: Cumulative TFP growth Decomposition with multiproduct term



Notes: This Figure decomposes the cumulative change in allocative efficiency in Figure 6 into single-product and multi-product terms.

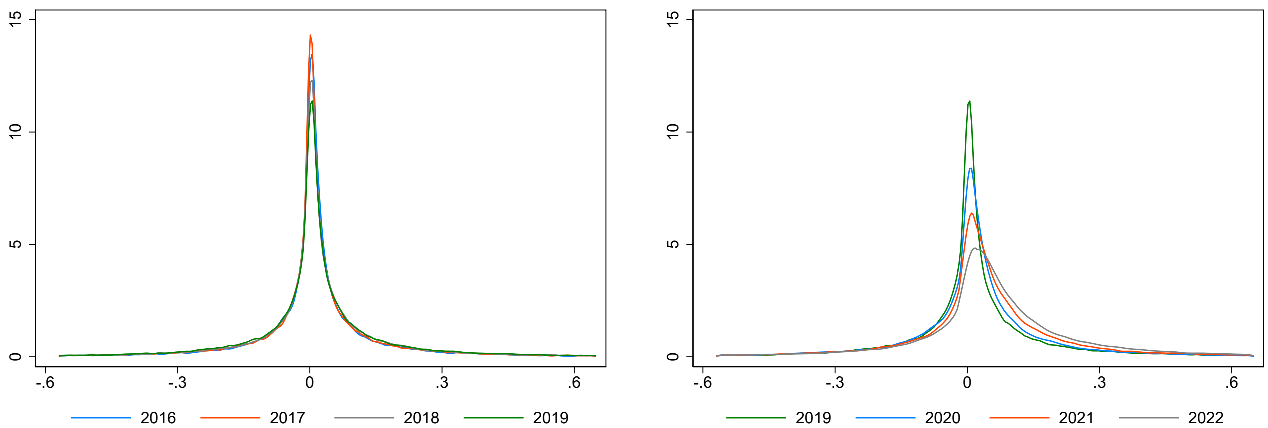
Furthermore, Figure 7 breaks down the allocative efficiency in Figure 6 into the multi-product and single product terms. It shows that the offsetting of reallocation due to joint production will become particularly strong after 2020. During this period, the economy was disrupted by COVID-19 and the subsequent high inflation. We interpret the latter as firms facing changes in product-specific demands, which change their total demand composition. As a response, firms were willing to readjust their product mix by reallocating productive resources. However, due to the non-infinite elasticity of transformation, firms were constrained to change their product mix.

Finally, the granularity of the data allows us to track the distributional changes of joint production (multiproduct term) limiting resource reallocation. Since the covariance degenerates to zero under the single-product firm assumption, the dispersion of covariance implies that joint-production forces are active. These distributions vary from period to period. Figure 8a plots the distribution for pre-COVID-19 (2016–2019), which is symmetric around 0, with slight differences from year to year.

Figure 8b presents the distribution after the onset of COVID-19, showing a shift to the right from year to year, resulting in a right-skewed distribution. This result suggests that the increase in the contribution from joint-production forces (multiproduct term) was not driven by a few specific firms.

Figure 8: $Cov_{s_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\gamma_i}\right)$ distributions by year

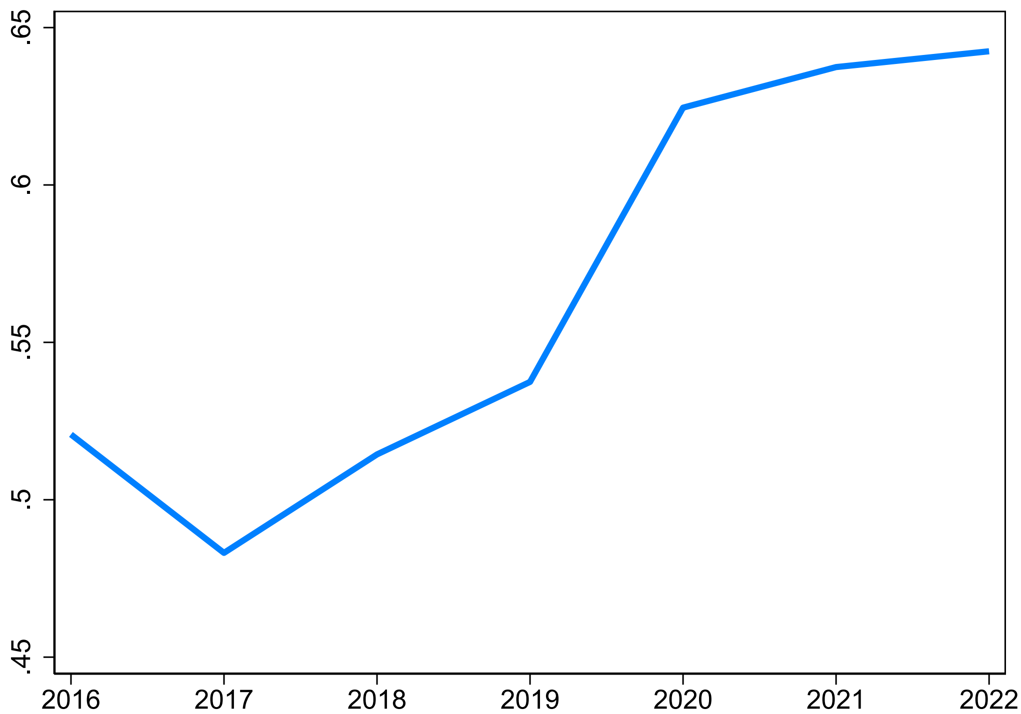
(a) 2016 - 2019 (b) 2019 - 2022



Notes: These Figures plot the distribution of firm level $Cov_{s_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\gamma_i}\right)$ for each year.

Finally, Figure 9 plots the median variance of product-specific production changes across firms from 2016 to 2022. This figure provides suggestive evidence that aligns with the changing distribution of multi-product firms shown in Figure 8b and corresponds to the period of significant contribution from the multi-product term in our decomposition. The increasing variance, particularly the sharp rise from 2019 to 2020 and its sustained high level thereafter indicates that firms have been under greater pressure to adjust their product portfolios. This trend coincides with the timeframe when we observe the most substantial impact of the multi-product term on allocative efficiency. The temporal consistency between the increased variance in product-specific production changes and the heightened contribution of the multi-product term reinforces our model's emphasis on the importance of multi-product firms engaged in joint production, especially during major economic shocks like the COVID-19 pandemic.

Figure 9: Large Product Mix Adjustments and Efficiency: Suggestive Evidence



Notes: This figure depicts the evolution of the median variance of product quantity changes, denoted as $Var_{\lambda_i}\left(d\log q_{ig}\right)$, from 2016 to 2022.

6 Extension: Ex-Ante Structural Results and Application

The empirical results presented in the previous Sections rely on price changes and factor reallocation in response to productivity and markup shocks, which are endogenously determined in general equilibrium. In this section, we develop a theoretical framework to model these endogenous responses and analyze how they depend on the economy's underlying structure. This approach extends our analysis beyond ex-post measurement to ex-ante prediction. Whereas our previous results relied on observing realized equilibrium outcomes, we can now use the model's structure to forecast responses to counterfactual shocks.

We apply this framework to study the "distance to the frontier," a widely used measure of resource misallocation. This method compares output in an efficient equilibrium (with all markup wedges removed) to that in a distorted decentralized economy. Our analysis demonstrates how the theoretical results of previous studies, such as Hsieh and Klenow (2009) and Baqaee and Farhi (2020), change when applying a joint production. Since an economy without markups is unobservable, a model is necessary to analyze this counterfactual case.

We begin by introducing the Nested CES-CET Model, which provides a tractable framework for our analysis. We then derive a linear system for price and sales responses, allowing us to characterize the economy's response to shocks. Using this system, we analyze the distance to the frontier, providing insights into the nature of misallocation in economies with joint production. Finally, we apply our framework to measure Chile's "distance to the frontier" under two assumptions—with and without joint-production—and assess its quantitative significance.

6.1 The Nested CES-CET Model

We introduce specific parametric assumptions into our general equilibrium system, specifying the nonparametric joint production function using the CES-CET parametric class. This approach offers both flexibility and tractability. The production function is given by:

$$\underbrace{\left(\sum_g \delta_{ig}\left[q_{ig}\right]^{\frac{\sigma_i-1}{\sigma_i}}\right)^{\frac{\sigma_i}{\sigma_i-1}}}_{\text{Output bundle}} = A_i \underbrace{\left(\sum_g \omega_{i,jg'} q_{i,jp}^{\frac{\theta_i-1}{\theta_i}} + \omega_L L_i^{\frac{\theta_i-1}{\theta_i}} + \omega_K K_i^{\frac{\theta_i-1}{\theta_i}}\right)^{\frac{\theta_i}{\theta_i-1}}}_{\text{Input Bundle}} \tag{7}$$

Here, σ_i represents the elasticity of transformation between different outputs, A_i denotes the productivity of firm i, and δ_{ig} are the share parameters for the outputs. The input bundle comprises intermediate inputs $q_{i,jg'}$, labor L_i, and capital K_i, aggregated using a CES function with an elasticity of substitution θ_i. For single-output firms, the production function simplifies to:

$$q_i = A_i \underbrace{\left(\sum_{j\in\mathcal{N},g'\in\mathcal{G}} \omega_{i,jg'} q_{i,jg'}^{\frac{\theta_i-1}{\theta_i}} + \omega_L L_i^{\frac{\theta_i-1}{\theta_i}} + \omega_K K_i^{\frac{\theta_i-1}{\theta_i}}\right)^{\frac{\theta_i}{\theta_i-1}}}_{\text{Input bundle}} \tag{8}$$

This class of models is highly general, nesting the nested CES system widely used in trade models and network macro models as a special case. For analytical simplicity, we assume a uniform substitution elasticity within the firm's CES structure, though extending the model to incorporate further nesting would be straightforward.

6.2 Linear System for Price and Sales Responses

Using this model, we construct a system for solving the first-order response to primitive shocks $(\boldsymbol{A}, \boldsymbol{\mu})$ to the endogenous variables. This system of linear equations, derived from the model's first-order conditions, enables us to generate ex-ante predictions of how the economy will respond to counterfactual shocks. We begin with the multi-product firm's forward equation under the CET output function:

Proposition 2. *Multi-Product firm's forward equation under CET output function:*

$$d\log p_{ig} = \underbrace{-\sum_{j\in\mathcal{N},g'\in\mathcal{G}} \tilde{\Psi}_{ig,jg'}\left(d\log\mu_j - d\log A_j\right) + \sum_{f\in\mathcal{F}} \tilde{\Psi}_{ig,f}\, d\log\Lambda_f}_{\textit{Indirect cost exposure}}$$
$$+ \underbrace{\sum_{j\in\mathcal{N},g'\in\mathcal{G}} c_j^R \tilde{\Psi}_{ig,jg'}\, d\log\Theta_{jg'}}_{\textit{Indirect exposure to the Product Mix adjustment}},$$

where

$$d\log\Theta_{jg'} = \left(\frac{\sigma_j}{\sigma_j+1}\right) d\log\mu_{jg'}/\mu_{jr} + \frac{1}{\left(1+\sigma_j\right)}\left[d\log\lambda_{jg'}/\lambda_{jr}\right],$$

and $c_j^R = \frac{mc_{jr}q_{jr}}{\sum mc_{jg}q_{jg}}$ *is a reference product cost share. Here, r denotes a reference good.*

This equation describes how changes in unit prices within a firm due to markups, productivity shocks, and price changes associated with endogenous product mix adjustments are transmitted through production networks to other firms' products. The first term illustrates the effect of exposure to common cost shocks on prices, a force present in standard production network models. The second term, unique to the joint production model, indicates the exposure of firms with nonlinear PPFs to changing product mix due to reallocation, which affects endogenous unit costs. The magnitude of this effect depends on the transformation elasticity, with lower elasticities leading to more significant cost effects. As σ approaches infinity, this endogenous effect disappears.

Next, we consider backward propagation:

Proposition 3. *Backward propagation:*

$$\lambda_{ig} d\log\lambda_{ig} = -\sum_{j\in\mathcal{N},g'\in\mathcal{G}} \lambda_{jg'}\left(\Psi_{jg',ig} - \mathbf{1}\left(ig = jg'\right)\right) d\log\mu_{jg'} + \sum_{k\in\mathcal{N},g''\in\mathcal{G}} \frac{\lambda_{kg''}}{\mu_{kg''}}\left(1-\theta^{k}\right) Cov_{\tilde{\Omega}_{(kg'',:)}}\left(d\log p, \Psi_{(:,ig)}\right).$$

This equation from Baqaee and Farhi (2020), describes the sales and factor share response. Notably, it does not contain σ, indicating that joint production only affects the variation of sales share through price. The equation shows how shocks propagate through the network, affecting markups and quantities of each product via changes in upstream suppliers' price indices and productivity.

By combining the forward equation from Proposition 1 and the backward equation from Proposition 2, we obtain a complete system of linear equations that characterizes the economy's response to shocks. This system consists of $2\times(\mathcal{N}\times\mathcal{G}+\mathcal{F})$ equations in $2\times(\mathcal{N}\times\mathcal{G}+\mathcal{F})$ unknowns, where N is the number of firms, G is the number of goods, and F is the number of factors. This system of equations fully characterizes the first-order response of all endogenous variables to any combination of productivity or markup shocks. By solving this linear system using standard matrix algebra, we can conduct counterfactual analyses and evaluate the impact of various shocks on the economy's equilibrium outcomes.

6.3 Distance to the frontier

Using our model, we can characterize the distance to the efficient frontier when introducing distortions, allowing us to predict efficiency losses from counterfactual changes in markups or other distortions:

Proposition 4. *Under joint production, starting at an efficient equilibrium, up to the second order, in response to the introduction of distortions, changes in the TFP are given by Domar-weighted Harberger triangles:*

$$\mathcal{L} = \frac{1}{2}\sum_{ig} \lambda_{ig} d\log q_{ig} d\log\mu_{ig}, \tag{9}$$

where λ_{ig} is the Domar weight of product g from firm i, q_{ig} is the quantity, and μ_{ig} is the markup.

This result demonstrates that TFP changes resulting from the introduction of distortions are solely determined by three statistics: Domar weights of each product, the magnitude of the wedges, and the change in the quantity of the product. The quantity change can be derived from sales and price changes given by our linear system, using the relationship $d\log q = d\log\lambda - d\log p$. Since this system includes transformation elasticity σ, it is generally affected by the elasticity value. To illustrate how joint production affects the Distance to the Frontier, we provide analytical solutions for two simple examples:

6.3.1 Horizontal Economy with Joint Production

We consider a horizontal economy similar to Hsieh and Klenow (2009), but with firms engaging in joint production of multiple goods. This allows us to investigate within-firm markup heterogeneity in the presence of production transformation constraints. In an economy with a representative consumer (CES utility with elasticity θ), N firms each use a common input L to produce G products using CET technology (elasticity σ). Markups μ_{ig} are heterogeneous across products and firms.

Proposition 5. *The distance to the frontier in the horizontal economy is given by:*

$$\mathcal{L} = -\frac{1}{2}\theta\left(\mathrm{Var}_{\boldsymbol{\lambda}}(d\log\mu_{ig}) - \frac{1}{\sigma+1}\mathbb{E}_{\bar{\lambda}}\left\{\mathrm{Var}_{s_i}(d\log\mu_{ig})\right\}\right),$$

where $\mathrm{Var}_{\boldsymbol{\lambda}}(d\log\mu_{ig})$ *and* $\mathrm{Var}_{\lambda i}(d\log\mu_{ig})$ *represent the Domar weighted variance of markup change and the average of within-firm variances of markup changes, respectively. The vectors* $\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{12}, \ldots, \lambda_{NG})$, $\bar{\boldsymbol{\lambda}} = (\lambda_1, \lambda_2, \ldots, \lambda_N)$ *is the vector of firm-level Domar weights with* $\lambda_i = \sum_g \lambda_{ig}$, *and* $\boldsymbol{s}_i = (\lambda_{i1}/\lambda_i, \lambda_{i2}/\lambda_i, \ldots, \lambda_{iG}/\lambda_i)$.

Proposition 5 characterizes the distance to the frontier in a horizontal economy with joint production and heterogeneous markups. The distance to the frontier denoted by L comprises two markup variances. The first term means the product-level markup's Domer-weighted variance gives the distance to the frontier. On the other hand, the second term is related to joint production. It means that the elasticity of finite transformation σ attenuates the effect of the variance of the markup within the firm. As σ increases and approaches infinity, the force of attenuation associated with joint production approaches zero. Conversely, as σ approaches zero (which implies Leontief production technology), the importance of within-firm markup dispersion decreases. We give this result in the following formal corollary.

Corollary 3 (Limit Cases)**.** *The distance to the frontier simplifies in extreme cases of the elasticity of transformation:*

1. *As* $\sigma \to \infty$ *(perfect substitution between products):*

$$\mathcal{L} = -\frac{1}{2}\theta \operatorname{Var}_{\lambda}(d \log \mu_{ig}).$$

2. *As* $\sigma \to 0$ *(Leontief production technology):*

$$\mathcal{L} = -\frac{1}{2}\theta \operatorname{Var}_{\bar{\lambda}}(\mathbb{E}_{s_i}(d \log \mu_i)).$$

In the case of perfect substitutes, misallocation depends on the variance of markups across all products. This term can be obtained by considering the product as an independent firm and applying the results of Baqaee and Farhi (2020). Conversely, in the Leontief case, only the variance of markups between firms is relevant. This is a consequence of the law of total variance. Given this result, one might be tempted to justify ignoring multi-product firms, as is common in much of the misallocation literature, especially when σ is small. However, this reasoning is specific to horizontal economies. This relationship easily breaks down in more complex economic structures that include inter-firm networks, and σ remains essential even when markups within firms are homogeneous. To illustrate this, we consider a simplified network example examined in Section 3.

6.3.2 Simplified Economy

Example 2. TFP Losses in a Simplified Economy: The second-order social loss is given by:

$$\mathcal{L} = -\frac{1}{4}\left(\frac{\sigma}{\sigma+1}\right)(d \log \mu)^2. \tag{10}$$

Despite zero within-firm markup dispersion, σ appears in the loss function, demonstrating that network structure and transformation elasticity jointly determine the distance to the frontier.

Welfare losses decrease as σ increases. We obtain an upper bound on social loss as σ approaches infinity. When σ approaches zero (Leontief case), no misallocation occurs regardless of markup size.

These findings emphasize the importance of considering network structures and transformation elasticities when evaluating misallocation and efficiency in economies with joint production, even without within-firm markup dispersion.

6.4 Application to Chile

TBD

7 Conclusion

This paper develops a theoretical framework to aggregate distortions in production networks with multiproduct firms. We assess their impact on aggregate TFP growth and derive a nonparametric sufficient statistic to describe allocative efficiency with multiproduct firms engaging in joint production.

We apply the framework to a comprehensive Chilean firm-to-firm transaction database. Reallocation effects considering joint production explain 86% of the observed aggregate TFP growth. Conversely, ignoring joint production leads to overestimating resource misallocation.

We demonstrate the importance of considering joint production in understanding aggregate TFP dynamics, especially during economic disruptions. The constraints multiproduct firms face in adjusting their product portfolios reduce reallocation within the network that single-product models would predict. This effect is pronounced in sectors such as mining and manufacturing, while services and construction emerge as potentially important sources of misallocation.

Our analysis reveals that joint production, a previously understudied source of TFP growth, can be of first-order importance. Our results demonstrate how aggregating granular microdata, through the lens of theory, can reduce the measure of our ignorance as captured by aggregate TFP and provide new insights into the drivers of economic growth.

References

Acevedo, P., Luttini, E., Pizarro, M., and Quevedo, D. (2023). Invoices rather than surveys: using ML to nominal and real indices.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica: journal of the Econometric Society*, 83(6):2411–2451.

Almunia, M., Antràs, P., Lopez-Rodriguez, D., and Morales, E. (2021). Venting out: Exports during a domestic slump. *The American economic review*, 111(11):3611–3662.

Bai, X., Fernández-Villaverde, J., Li, Y., and Zanetti, F. (2024). The causal effects of global supply chain disruptions on macroeconomic outcomes: Evidence and theory. *SSRN Electronic Journal*.

Baqaee, D., Burstein, A., Duprez, C., and Farhi, E. (2023). *Supplier Churn and Growth: A Micro-to-Macro Analysis*. National Bureau of Economic Research.

Baqaee, D. R. and Farhi, E. (2020). Productivity and misallocation in general equilibrium. *The quarterly journal of economics*, 135(1):105–163.

Basu, S. and Fernald, J. G. (2002). Aggregate productivity and aggregate technology. *European economic review*, 46(6):963–991.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2010). Multiple-product firms and product switching. *The American economic review*, 100(1):70–97.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2011). Multiproduct firms and trade liberalization. *The Quarterly Journal of Economics*, 126(3):1271–1318.

Bigio, S. and La'O, J. (2020). Distortions in production networks*. *The quarterly journal of economics*, 135(4):2187–2253.

Boehm, C. E., Flaaen, A., and Pandalai-Nayar, N. (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake. *The review of economics and statistics*, 101(1):60–75.

Boehm, J. and Oberfield, E. (2023). Growth and the fragmentation of production.

Burstein, A., Cravino, J., and Rojas, M. (2024). Input price dispersion across buyers and misallocation.

Cairncross, J. and Morrow, P. (2023). Multi-product markups.

Carrillo, P., Donaldson, D., Pomeranz, D., and Singhal, M. (2023). *Misallocation in Firm Production: A Nonparametric Analysis Using Procurement Lotteries*. National Bureau of Economic Research.

Carvalho, V. M., Nirei, M., Saito, Y. U., and Tahbaz-Salehi, A. (2020). Supply chain disruptions: Evidence from the great east japan earthquake*. *The quarterly journal of economics*, 136(2):1255–1321.

CNEP, C. N. E. . P. C. (2023). Informe anual de productividad 2023.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica: journal of the Econometric Society*, 84(2):445–510.

Dhyne, E., Petrin, A., Smeets, V., and Warzynski, F. (2017). Multi product firms, import competition, and the evolution of firm-product technical efficiencies. (23637).

Dhyne, E., Petrin, A., Smeets, V., and Warzynski, F. (2022). Theory for extending single-product production function estimation to multi-product settings. (30784).

Diewert, W. E. (1971). An application of the shephard duality theorem: A generalized leontief production function. *The journal of political economy*, 79(3):481–507.

Diewert, W. E. (1973). Functional forms for profit and transformation functions. *Journal of Economic Theory*.

Ding, X. (2023). *Industry Linkages from Joint Production*. U.S. Census Bureau, Center for Economic Studies.

Fujiy, B. C., Ghose, D., and Khanna, G. (2022). Production networks and firm-level elasticities of substitution.

Giroud, X. and Mueller, H. M. (2019). Firms' internal networks and local economic shocks. *The American economic review*, 109(10):3617–3649.

Hall, R. (1988). The relation between price and marginal cost in U.S. industry. *The journal of political economy*, 96(5):921–947.

Hall, R. E. (1973). The specification of technology with several kinds of output. *The journal of political economy*, 81(4):878–892.

Hottman, C. J., Redding, S. J., and Weinstein, D. E. (2016). Quantifying the sources of firm heterogeneity. *The Quarterly Journal of Economics*, 131(3):1291–1364.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in china and india. *The quarterly journal of economics*, 124(4):1403–1448.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *The Review of economic studies*, 45(3):511–518.

Kikkawa, A. K. (2022). Imperfect competition in firm-to-firm trade. *Journal of the European Economic Association*, 20(5):1933–1970.

Kim, R. (2020). The effect of the credit crunch on output price dynamics: The corporate inventory and liquidity management channel. *The Quarterly Journal of Economics*, 136(1):563–619.

Klette, T. J. and Kortum, S. (2004). Innovating firms and aggregate innovation. *The journal of political economy*, 112(5):986–1018.

Lau, L. J. (1972). Profit functions of technologies with multiple inputs and outputs. *The review of economics and statistics*, 54(3):281–289.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. P. (2014). Market size, competition, and the product mix of exporters. *The American economic review*, 104(2):495–536.

Mayer, T., Melitz, M. J., and Ottaviano, G. I. P. (2021). Product mix and firm productivity responses to trade competition. *The review of economics and statistics*, 103(5):874–891.

Osotimehin, S. and Popov, L. (2023). Misallocation and intersectoral linkages. *Review of economic dynamics*.

Petrin, A. and Levinsohn, J. (2012). Measuring aggregate productivity growth using plant-level data. *The Rand journal of economics*, 43(4):705–725.

Powell, A. A. and Gruen, F. H. G. (1968). The constant elasticity of transformation production frontier and linear supply system. *International economic review*, 9(3):315–328.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of economic dynamics*, 11(4):707–720.

Solow, R. M. (1957). Technical change and the aggregate production function.

Valmari, N. (2023). Estimating production functions of multiproduct firms. *The Review of economic studies*.

Appendix

A Additional Figures and Tables

A.1 Additional Empirical Results for Reduced-Form Evidence

In this appendix section, we present additional Figures and Table for the event study analysis shown in Section 2.

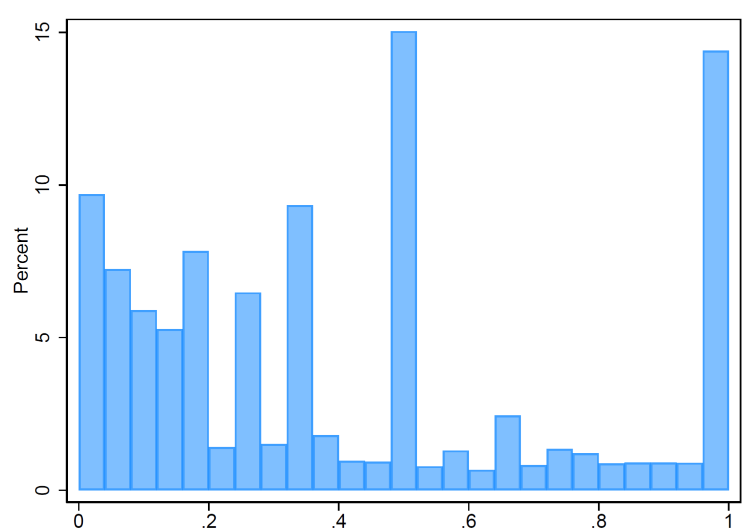
Firms Sell Different Products to Distinct Sets of Buyers

First, we construct the following measure to characterize the heterogeneity from the intermediate inputs buyer perspective across products and within firms:

$$S_i = \frac{\text{number of buyers of the main product of firm } i}{\text{number of buyers of firm } i},$$

where the main product is that with the largest sales within firm i in 2018. Figure A1 presents the distribution of this measure across firms.

Figure A1: Buyer heterogeneity across product within firm



Notes: Histogram of the number of buyers buying the main product of firm i / number of buyers in firm i for a multiproduct firm. The main product is the product with the highest sales within that firm. Data are from 2018.

If the buyers of the seller's main product and its other products were the same, S_i would be one. Some mass exists at $S_i = 1$ but for more than 50% of multiproduct firms; however, buyers of the main product constitute less than 50% of the total buyer-firms base. The fact that each product has a distinct set of buyers ensures that we can construct a sample where the main product experiences a demand shock while the other products do not.

Characteristics of Treatment Firms

Next, Table A1 displays the characteristics of treated and control firms.

Table A1: Characteristics of Treatment Firms

	Treatment Firms	Control Firms
Number of firms	26,411	96,321
Number of workers	6	4
Number of products sold	16	10
Annual revenue (million pesos)	186	101
Annual total intermediate purchases (million pesos)	107	59
Share of firms in manufacturing	0.21	0.24
Share of firms in Retail and wholesale	0.44	0.39
Share of firms in Services	0.22	0.21

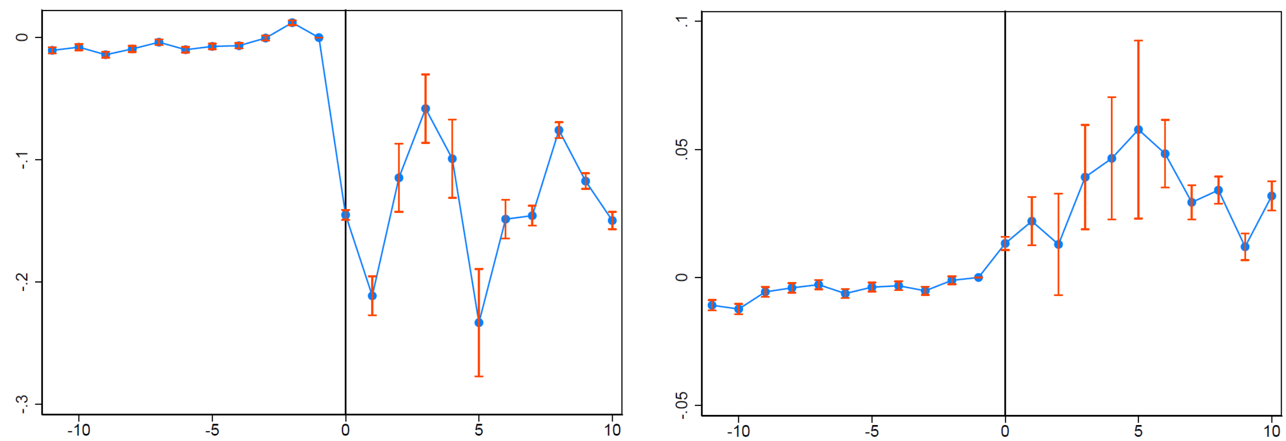
Notes: This table presents the characteristics of treated firms (those whose major product buyers experienced lockdowns in March 2020) and control firms, showing values from February 2020, the month before the shock. The rows for the number of workers, the number of products sold, revenue, and total intermediate purchases display the median of each statistic. The industry shares indicate the proportion of firms within each group belonging to specific industries.

Event study excluding small firms

To address potential bias from financially constrained firms in our event study, we conduct a robustness check focusing on larger firms, which we assume to be relatively less financially constrained. We calculate total firm sales, including both network and final consumer sales, for each year of the study period (2019-2021). We then isolate firms above the 90th percentile of this distribution and replicate our event study analysis from Section 2 using only this subset of larger firms. The results of this robustness check are presented in Figure A2.

Figure A2: The effects of demand shocks to the main product on the production of other products within the firm: Robustness check

(a) Log Quantity (b) Log Price



Notes: Standard errors are clustered at the firm-county level, and the error bands represent 95% confidence intervals. The X-axis represents the time to treat, with 0 denoting March 2020 when the main product experienced the demand shock, and other values indicating the number of months before or after this event.

The event study results excluding small firms are similar to the event study including all firms, suggesting that financial constraints are not driving the results.

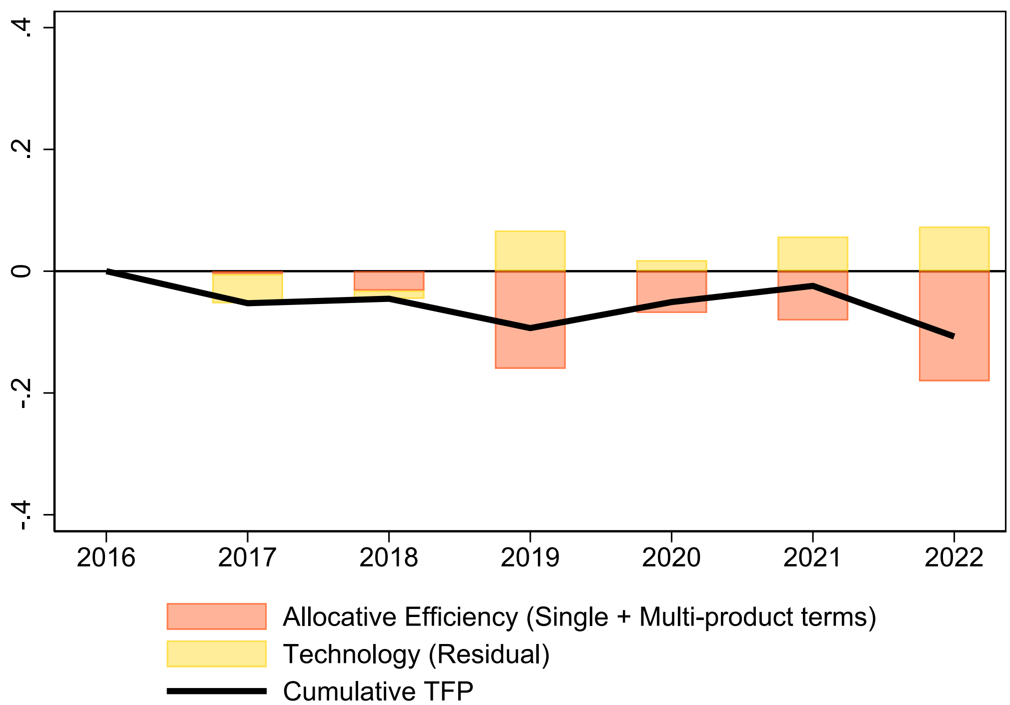
A.2 Additional Empirical Results for Growth accounting

Growth accounting results assuming that markups within firms are equal in initial equilibrium

We show the results when the equilibrium is approximated around $\mu_{ig} = \mu_i$. Note that the level of μ_i does not affect the results since Γ is normalized for covariance.

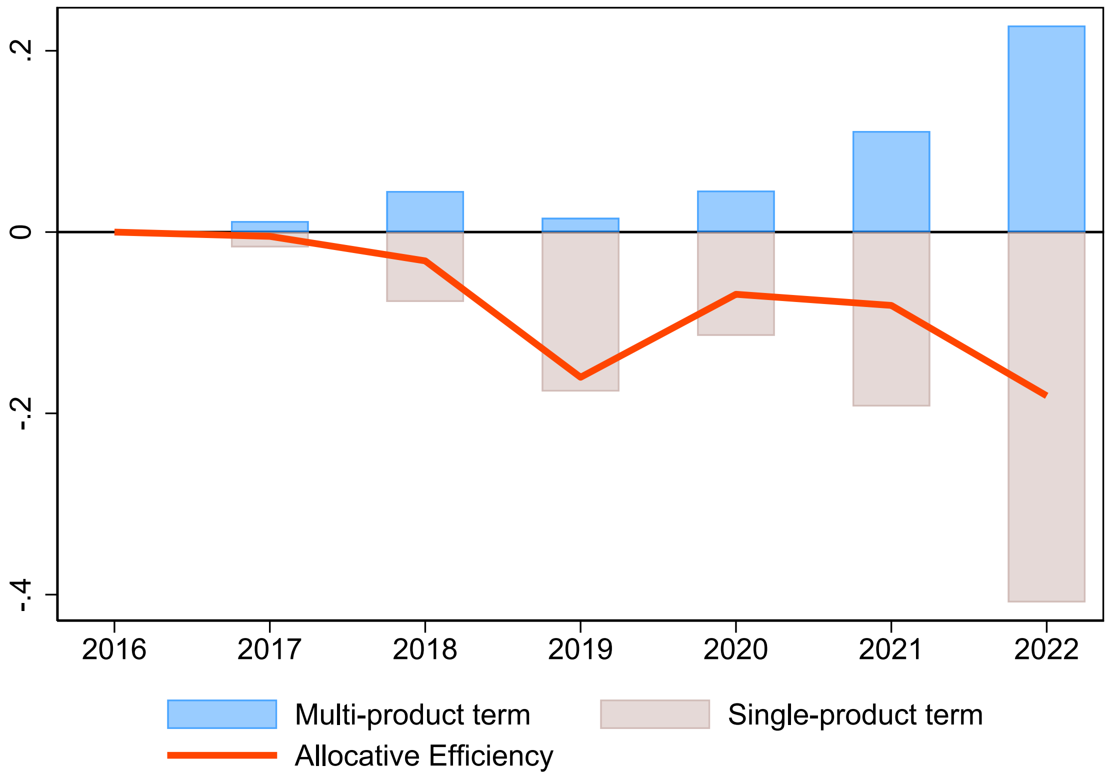
$$\Gamma_{ig} \equiv \underbrace{\frac{\tilde{\lambda}_{ig}}{salesshare_{ig}}}_{\text{Downstream distortion}} \times \underbrace{\mu_i}_{\text{own markup}},$$

Figure A3: Cumulative TFP Decomposition with Multi-Product Term



Notes: This Figure shows the cumulative log change calculated by repeatedly applying the equation from Proposition 1 each year. Technology (residual) is calculated by subtracting allocative efficiency from TFP.

Figure A4: Cumulative TFP growth Decomposition with multiproduct term

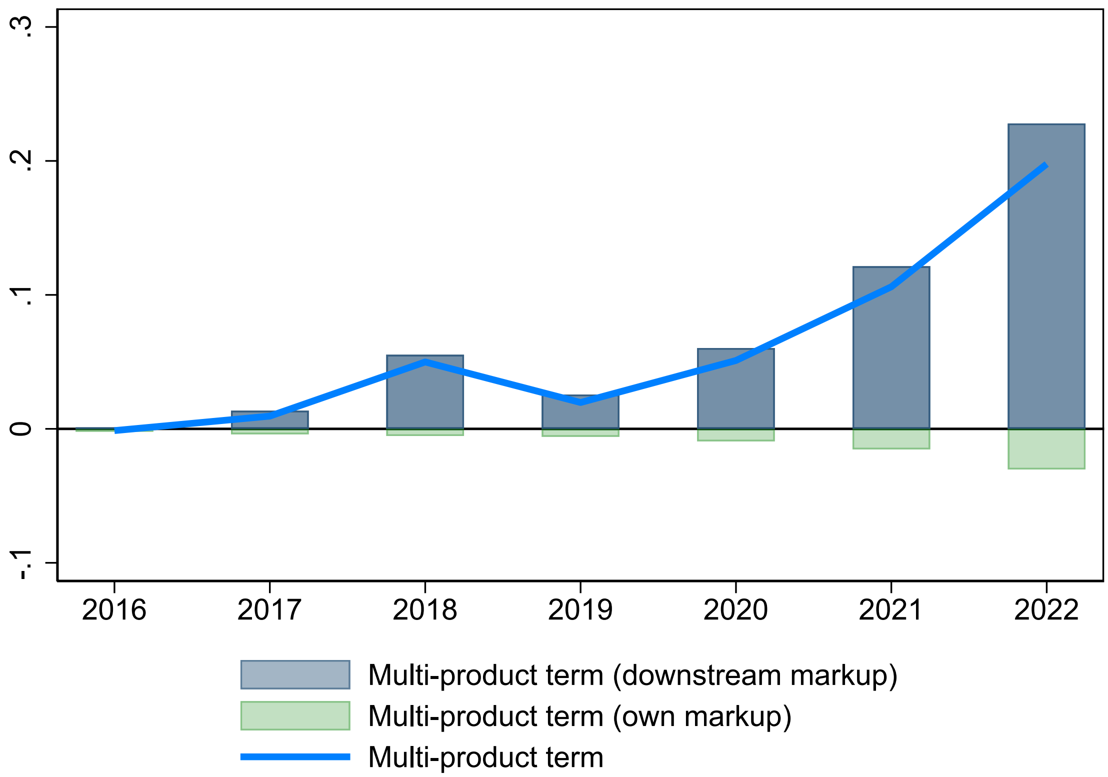


Notes: This Figure decomposes the cumulative change in allocative efficiency in Figure A3 into single-product and multi-product terms.

Decomposition of Allocative Efficiency

Decompose whether the variation in the covariance of allocative efficiency is due to one's own markup or downstream markup.

Figure A5: Cumulative TFP growth Decomposition with multi-product term



Notes: This Figure shows the cumulative log change calculated by repeatedly applying the equation from Proposition 1 each year.

Table A2: Aggregate firm-level statistics

Year	Count	Sales	Wagebill	Employment
2016	110,451	262,506	40,260	4,242,555
2017	114,480	277,960	43,691	4,349,248
2018	115,916	330,486	44,688	4,349,454
2019	116,706	336,386	47,299	4,425,780
2020	102,306	310,317	44,053	3,935,883
2021	105,651	376,220	51,642	4,166,838
2022	105,032	454,818	59,148	4,266,972

Notes: Count stands by the number of firms, while sales and wage bill are yearly aggregates expressed in million pesos. Employment represents the headcount of yearly workers included in the sample.

B Downstream Markup

Table A3: The 30 most distorted products

Ranking	Description
1	Insurance brokerage services
2	Other services
3	Passenger air transport services
4	Wholesale trade intermediary services
5	Electricity distribution and other related services
6	Investigation and security services
7	Airport services
8	Radio and open TV broadcast services
9	Wastewater treatment services
10	Online content services
11	Cleaning services
12	Liquefied Natural Gas
13	Employment services (placement and supply)
14	Postal and courier services
15	Tobacco
16	Paper and cardboard containers, paper or cardboard for recycling
17	Other IT services
18	News agency services
19	Margarine and similar preparations, other residues and waste from fats
20	General insurance
21	Other rubber products
22	Other auxiliary and complementary services for education services
23	Other goods or services not classified elsewhere
24	Long-distance passenger transport services
25	Gas distribution services and other related services
26	Some other product
27	Maritime passenger transport services
28	Research and development services
29	Repair and installation of machinery and equipment, except for the textile industry
30	Database software licensing services

Notes: For 2018, products are ranked using the network distortion medians for CUP's product categories, and products with the top 30 downstream markup sizes are reported.

Table A4: The 30 Least distorted product

Ranking	Description
1	Molybdenum minerals and their concentrates
2	Other non-metallic minerals
3	Gaseous natural gas
4	Crude oil
5	Mining works
6	Unrefined copper, ashes, residues and wastes of copper
7	Silver
8	Public administration and defense services; compulsory social security plans
9	Pet food
10	Bird food
11	Fish meal, crustacean, mollusk and other aquatic invertebrate meal
12	Ammonium nitrate
13	Lease services with or without purchase option
14	Bread
15	Veterinary services
16	Poultry meat and edible offal
17	Integrated telecommunications services (packs)
18	Fuel oil
19	Beers
20	Life insurance
21	Cakes, cakes and cookies
22	Hake
23	Consultancy and post services
24	Copper minerals and their concentrates
25	Public hospital services
26	Social and association services
27	Petroleum gas and other gaseous hydrocarbons, except natural gas
28	Fish oil
29	Mining exploration and evaluation services
30	Housing services

Notes: For 2018, products are ranked using the network distortion medians for CUP's product categories, and products with the top 30 downstream markup sizes are reported.

C Detailed Methodology for Product-Level Markup Estimation

We estimate product-level markups following the production approach based on Dhyne et al. (2022). In a joint production setup, firms use common inputs to produce a product portfolio, meaning that some inputs may simultaneously be used to produce multiple products. They proposed an Ackerberg et al. (2015) like production function estimation method based on Diewert (1973)'s production set. The following is an overview of Dhyne et al. (2022)'s methodology.

A firm has production possibilities set, V, that consists of a set of feasible inputs $\boldsymbol{x} = (x_1, ..., x_M)$ and outputs of the product,$\boldsymbol{q} = (q_1, ..., q_G)$. For any (q_g, x) the transformation function is defined as

$$q_g^* = f_g\left(q_g, x\right) \equiv \max\left\{q_g | \left(q_g, q_{-g}, x\right) \in V\right\}$$

To identify the unobserved marginal cost for each firm's product, we rely on (variable) cost minimization. Firms have $N-1$ freely variable inputs and one fixed input, capital (K), so the problem that a firm faces to minimize its variables cost to produce its output vector q_i^* given the input prices vector $\boldsymbol{p_x} = (p_{x1}, ..., p_{xM})$ and unobserved productivity for products, $\boldsymbol{\omega} = (\omega_1, ..., \omega_G)$.

Defining the Lagrangian multiplier of the cost minimization problem, mc_g, as the marginal cost of product g, the first order condition for every optimal input demand yield:

$$p_m = mc_g \frac{\partial f(q_{-g}^*, x, K, \boldsymbol{\omega})}{\partial x_m} \quad \forall m = 1, .., M,$$

It is possible to solve for product g marginal cost as the expenditure on production input m divided by its output elasticity (β_n^g) times product g production:

$$mc_g = \frac{p_m}{\frac{\partial f(q_{-g}^*, x, K, \omega)}{\partial x_m}} = \frac{p_m x_m^*}{\beta_m^g q_g^*},$$

Multiplying the marginal cost expression by $\frac{1}{p_g}$, where p_g is product g price, product g markup is given by:

$$\mu_g = \beta_m^g \frac{p_g q_g^*}{p_m x_m^*},$$

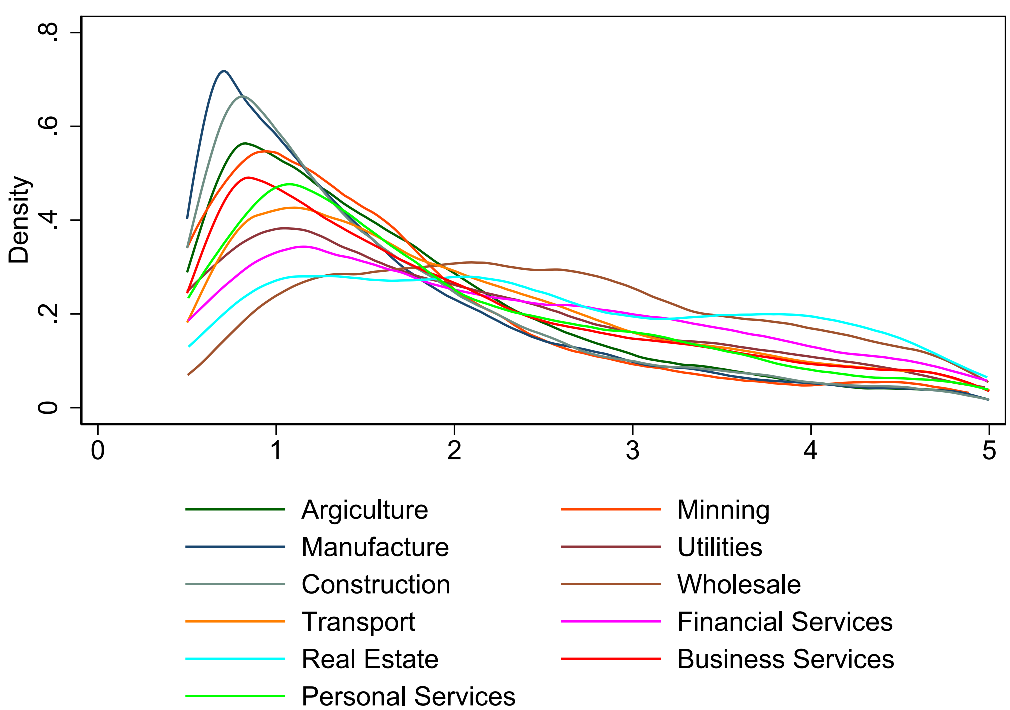
We use control functions for the unobserved productivity terms (i.e., Ackerberg et al. (2015)) to account for unobserved productivity with the difference of the need for instruments for q_{-g}; following Dhyne et al. (2022) we use lagged values of q_{-g}. We assume that firms use a Cobb-Douglas production function with three factors: (Capital K, Labor L, and Materials M). A multi-product firm will produce physical units of product g using the following production function:

$$\log q_{gt} = \beta_0^g + \beta_k^g \log k_t + \beta_l^g \log l_t + \beta_m^g \log m_t^j + \gamma_{-g}^g \log q_{-gt} + \omega_{gt}$$

We pool together products at one digit (12 aggregate product categories) and perform production function estimations separately for each category following ACF using a GMM estimator.

Product-level markup distribution concentrated around 1, with a 1.22 median. We remain agnostic about product-level markup interpretation.

Figure A6: Product-level markup distribution in 2018



1

D Proofs

Proof of Example 1 in subsection 3.6:

Proof. First, we compute relative network distortions:

$$\gamma_{ig} \equiv \frac{\Gamma_{ig}}{\Gamma_i},$$

where the average distortion of firm i is defined as: $\Gamma_i \equiv \sum_g \tilde{\lambda}_{ig} / \sum_g \left(salesshare_{ig} / \mu_{ig} \right)$. So Γ_1 is

$$\Gamma_1 \equiv \frac{1}{\frac{1}{\mu} + \frac{1}{\mu^2}} 2 = \frac{2}{\frac{1}{\mu}\left(1 + \frac{1}{\mu}\right)}.$$

and

$$\gamma_{11} = \frac{\mu^2}{\frac{2}{\frac{1}{\mu}\left(1+\frac{1}{\mu}\right)}} = \frac{(\mu + 1)}{2},$$

$$\gamma_{12} = \frac{\mu}{\frac{2}{\frac{1}{\mu}\left(1+\frac{1}{\mu}\right)}} = \frac{\mu \frac{1}{\mu}\left(1 + \frac{1}{\mu}\right)}{2} = \frac{\left(1 + \frac{1}{\mu}\right)}{2},$$

Thus,

$$\frac{1}{\gamma_{11}} = \frac{2}{\frac{1+\mu}{\mu}} = 2\frac{1}{(\mu + 1)},$$

$$\frac{1}{\gamma_{12}} = \frac{2}{\frac{1+\mu}{\mu}} = 2\frac{\mu}{1 + \mu}.$$

Substitute them into covariance:

$$Cov_{s_i}\left(d \log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i}\right) = Cov_{\left[\frac{1}{2},\frac{1}{2}\right]}\left(\begin{bmatrix} -\epsilon \\ 0 \end{bmatrix}, \begin{bmatrix} 2\frac{1}{(\mu+1)} \\ 2\frac{\mu}{(\mu+1)} \end{bmatrix}\right),$$
$$= \left(\frac{\mu - 1}{2\,(1 + \mu)}\right)\epsilon.$$

Next, we compute the endogenous response of Labor share $d \log \Lambda$ to markup shock. By factor share identity, we know

$$\Lambda_L = 1 - \frac{1}{2}\left(1 - \frac{1}{\mu_{12}}\right) - \frac{1}{2}\left(1 - \frac{1}{\mu_{21}}\right) - \frac{1}{2}\frac{1}{\mu_{21}}\left(1 - \frac{1}{\mu_{11}}\right),$$

Evaluating markup other than μ_{21} with μ and take the logarithm, we have

$$\log \Lambda = \log \frac{1}{\mu} + \log \frac{1}{2}\left(1 + \frac{1}{\mu_{21}}\right).$$

Taking the derivative with respect to $\log \mu_{21}$,

$$\begin{aligned}
\frac{d \log \Lambda}{d \log \mu_{21}} &= \frac{d \log\left(\frac{1}{\mu}\frac{1}{2}\left(1 + \frac{1}{\mu_{21}}\right)\right)}{d\mu_{21}} \frac{d\mu_{21}}{d \log \mu_{21}}, \\
&= \frac{-\frac{1}{\mu_{21}}}{\frac{1}{2}\left(1 + \frac{1}{\mu_{21}}\right)}, \\
&= \frac{-\frac{1}{\mu_{21}}}{\frac{1+\mu}{\mu}}.
\end{aligned}$$

Evaluate at $\mu_{11} = \mu$ gives

$$\frac{d \log \Lambda}{d \log \mu_{21}} = -\frac{1}{\mu + 1}.$$

And

$$\sum_i \tilde{\lambda}_i d \log \mu_i = \frac{1}{2}(-\epsilon).$$

Therefore,

$$-d \log \Lambda - \sum_i \tilde{\lambda}_i d \log \mu_i = \left(\frac{\mu - 1}{2\,(1 + \mu)}\right)\epsilon.$$

□

Proof of Example 2 in subsection 3.6:

Proof. Pick product 2 to be a reference product for firm 1. Then, we know

$$d \log p_{11}/p_{12} = d \log \mu_{11}/\mu_{12} + \frac{1}{\sigma} d \log y_{11}/y_{12}.$$

Using $d \log \lambda = d \log p + d \log y$

$$d \log p_{11}/p_{12} = \left(\frac{\sigma}{\sigma + 1}\right) d \log \mu_{11}/\mu_{12} + \frac{1}{\sigma + 1} d \log \lambda_{11}/\lambda_{12},$$

where $d \log \mu_{11} = 0$, $d \log \mu_{12} = 0$, and $d \log \mu_{21} = 0$. By Cobb Douglas assumption, we know $d\log\lambda_{11}/\lambda_{12} = -\epsilon$. Thus we have

$$d\log p_{11}/p_{12} = -\left(\frac{1}{\sigma+1}\right)\epsilon.$$

$$\begin{aligned}
d\log TFP &= Cov_{[\frac{1}{2},\frac{1}{2}]}\left(\begin{bmatrix} d\log p_{11} \\ d\log p_{12} \end{bmatrix}, \begin{bmatrix} \frac{1}{\gamma_1} \\ \frac{1}{\gamma_2} \end{bmatrix}\right) - d\log\Lambda - \sum_i \tilde{\lambda}_i d\log\mu_i, \\
&= Cov_{[\frac{1}{2},\frac{1}{2}]}\left(\begin{bmatrix} d\log p_{11}/p_{12} \\ d\log p_{12}/p_{12} \end{bmatrix}, \begin{bmatrix} 2\frac{\mu}{1+\mu} \\ 2\frac{1}{(\mu+1)} \end{bmatrix}\right) + \left(\frac{\mu-1}{2(1+\mu)}\right)\epsilon, \\
&= Cov_{[\frac{1}{2},\frac{1}{2}]}\left(\begin{bmatrix} -\left(\frac{1}{\sigma+1}\right)\epsilon \\ 0 \end{bmatrix}, \begin{bmatrix} 2\frac{\mu}{1+\mu} \\ 2\frac{1}{(\mu+1)} \end{bmatrix}\right) + \left(\frac{\mu-1}{2(1+\mu)}\right)\epsilon, \\
&= \left(\frac{\sigma}{\sigma+1}\right)\left(\frac{\mu-1}{2(1+\mu)}\right)\epsilon.
\end{aligned}$$

□

Proof of Proposition 1

Lemma 1. *Price equation with multi-product firms for some reference product r of firm i:*

$$\frac{y_{ir}mc_{ir}}{C(y_i,p)}d\log p_{ir} = -d\log A_i/\mu_i + \underbrace{\sum_{j,k}\frac{p_{jg}x_{i,jg'}}{C(y_i,p)}d\log p_{jg'} + \sum_f \frac{w_f l_{if}}{C(y_i,p)}d\log w_f}_{\text{intermediate and factorprice}}$$

$$+ \underbrace{\sum_{g\neq r}\left(-\frac{y_{ig}mc_{ig}}{C(y_i,p)}\right)d\log p_{ig}}_{\text{other product from the same firm}},$$

Proof. By CRS, we know

$$C_i(\boldsymbol{q}_i,\boldsymbol{p}_i) = \sum_g q_{ig}mc_{ig}.$$

Total derivative:

$$
\begin{aligned}
RHS &= \sum_{g} \frac{\partial \log\left(\sum q_{ig} mc_{ig}\right)}{\partial \log y_i} y_i d\log y_i + \sum_{i} \frac{\partial \log\left(\sum q_{ig} mc_{ig}\right)}{\partial \log mc_{ig}} mc_{ig} d\log mc_{ig}, \\
&= \sum_{g} \frac{q_{ig} mc_{ig}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log q_{ig} + \sum_{g} \frac{q_{ig} mc_{ig}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log mc_{ig}.
\end{aligned}
$$

and

$$
LHS = -d\log A_i + \sum_{j,g'} \frac{p_{jg'} x_{i,jg'}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log p_{jg'} + \sum_{f} \frac{w_f l_{i,f}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log w_f + \sum_{g} \frac{q_{ig} mc_{ig}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log q_{ig}.
$$

Hence,

$$
\sum_{g} \frac{y_{ig} mc_{ig}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log mc_{ig} = -d\log A_i + \sum_{j,g'} \frac{p_{jg'} x_{i,jg'}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log p_{jg'} + \sum_{f} \frac{w_f L_{if}}{C_i\left(\boldsymbol{q}_i, \boldsymbol{p}_i\right)} d\log w_f. \tag{11}
$$

□

Cost minimization with joint production function:

Pick some reference product r of firm i. Following Hall (1973), as a concequence of cost minimization, the following condition holds:

$$
\frac{mc_{ig}}{mc_{ir}} = \frac{\partial F_i^Q\left(\boldsymbol{q}\right)/\partial q_{ig}}{\partial F_i^Q\left(\boldsymbol{q}\right)/\partial q_{ir}},
$$

By taking the log difference and adjusting it with markup, we get the following formula:

$$
d\log\left(p_{ir}/p_{ig}\right) = d\log\left(\mu_{ir}/\mu_{ig}\right) + d\log\left(\frac{\partial F^Q\left(\boldsymbol{q}\right)/\partial q_{ig}}{\partial F^Q\left(\boldsymbol{q}\right)/\partial q_{ir}}\right). \tag{12}
$$

This pins down the equilibrium prices with equation 11. For later proof, we define the RHS of the equation 12 as Θ_{ig}.

$$
d\log p_{ig}/p_{ir} = \Theta_{ig}. \tag{13}
$$

Then, we proceed to the main proof.

Proof. From Lemma 1 We know for one reference product r:

$$\frac{q_{ir}mc_{ir}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log mc_{ir} = -\,d\log A_i + \sum_{j,p}\frac{p_{jp}x_{jp}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{jp} + \sum_f \frac{w_f L_{if}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log w_f + \sum_{g\neq r}\left(-\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)d\log mc_{ig},$$

$$\begin{aligned}\frac{q_{ir}mc_{ir}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{ir} &= -\,d\log A_i + d\log\mu_i + \sum_{j,p}\frac{p_{jp}x_{i,jp}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{jp} + \sum_f \frac{w_f L_{if}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log w_f \\ &\quad + \sum_{g\neq r}\left(-\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)d\log p_{ig}.\end{aligned} \tag{14}$$

From equation 13, we have

$$d\log p_{ig}/p_{ir} = \Theta_{ig}.$$

Combining 11 with 13 yields

$$\begin{aligned}\frac{q_{ir}mc_{ir}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{ir} &= -d\log A_i + \sum_g \frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log\mu_{ig} + \sum_{j,p}\frac{p_{jp}x_{jp}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{jp} + \sum_f \frac{w_f l_{if}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log w_f \\ &\quad + \sum\left(-\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)d\log p_{ig}, \\ \frac{q_{ir}mc_{ir}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}d\log p_{ir} &= -d\log A_i + d\log\mu_i + \sum_{j,p}\tilde{\Omega}_{ig,jp}d\log p_{jp} + \sum_f \tilde{\Omega}_{ig,f}d\log w_f + \sum_{g\neq r}\left(-\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)\left[d\log p_{ir} + \Theta_{ig}\right], \\ d\log p_{ir} &= -d\log A_i + d\log\mu_i + \sum_{j,p}\tilde{\Omega}_{ig,jp}d\log p_{jp} + \sum_f \tilde{\Omega}_{ig,f}d\log w_f + \sum_{g\neq r}\left(-\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)\Theta_{ig}.\end{aligned}$$

Since $\Theta_{ig} = 0$ if g is an reference products, the price equations could be written by

$$d\log p_{ig} = -d\log A_i + d\log\mu_i + \sum_{j,g'}\tilde{\Omega}_{ig,jg'}d\log p_{jg'} + \sum_f \tilde{\Omega}_{ig,f}d\log w_f + \left\{\mathbb{I}_i(g) - \sum_{g\neq r}\left(\frac{q_{ig}mc_{ig}}{C_i\left(\boldsymbol{q}_i,\boldsymbol{p}_i\right)}\right)\right\}\Theta_{ig}.$$

In vector notation

$$d\log\boldsymbol{p} = (I - \tilde{\Omega}^{\mathcal{NG}\times\mathcal{NG}})^{-1}\left\{-d\log\boldsymbol{A}^{\mathcal{NG}\times 1} + d\log\boldsymbol{\mu}^{\mathcal{NG}\times 1} + \tilde{\Omega}_f^{\mathcal{NG}\times\mathcal{F}}\,d\log\boldsymbol{w} + (\mathbf{1} - \mathbf{C})\circ\boldsymbol{\Theta}^{\mathcal{NG}\times 1}\right\},$$

where $\circ$ represents Hadamard product and and C is a vector of $\mathcal{NG}\times 1$, with the following C_i common elements for firm $i \in \mathcal{N}$,

$$C_i = \sum_{g \neq r} \left(\frac{q_{ig} mc_{ig}}{C(\boldsymbol{q}_i, \boldsymbol{p}_i)} \right).$$

We know

$$d \log Y = -\boldsymbol{b}' d \log \boldsymbol{p}.$$

$$\begin{aligned} d \log Y &= -\boldsymbol{b}' \tilde{\Psi}^{\mathcal{NG} \times \mathcal{NG}} \left\{ -d \log \boldsymbol{A} + d \log \boldsymbol{\mu} + \tilde{\Omega}_f d \log \boldsymbol{w} + (\mathbf{1} - \mathbf{C}) \circ \boldsymbol{\Theta}^{\mathcal{NG} \times 1} \right\}, \\ &= -\tilde{\boldsymbol{\lambda}}' \left\{ -d \log \boldsymbol{A} + d \log \boldsymbol{\mu} + \tilde{\Omega}_f d \log \boldsymbol{w} + (\mathbf{1} - \mathbf{C}) \circ \boldsymbol{\Theta}^{\mathcal{NG} \times 1} \right\}. \end{aligned}$$

subtracting $\sum_f \tilde{\Lambda}_f d \log L_f$ from both sides yields

$$\begin{aligned} d \log TFP = & \sum_{i \in \mathcal{N}} \tilde{\lambda}_i d \log A_i - \sum_{i \in \mathcal{N}} \tilde{\lambda}_i d \log \mu_i - \sum_{f \in \mathcal{F}} \tilde{\Lambda}_f d \log \Lambda_f, \\ & - \sum_i \left(\sum_{g \in \mathcal{G}} \tilde{\lambda}_{ig} d \log p_{ig}/p_{ir} - \sum_{g \neq r} \frac{q_{ig} mc_{ig}}{C_i(\boldsymbol{q}_i, \boldsymbol{p}_i)} \tilde{\lambda}_i d \log p_{ig}/p_{ir} \right). \end{aligned}$$

$$\begin{aligned} \left(\sum_g \tilde{\lambda}_{ig} d \log p_{ig}/p_{ir} - \sum_{g \neq r} \frac{q_{ig} mc_{ig}}{C_i(y, p)} \tilde{\lambda}_i d \log p_{ig}/p_{ir} \right) &= \tilde{\lambda}_i \left(\sum_{g \in \mathcal{G}} s_{ig} d \log p_{ig}/p_{ir} - \sum_{g \neq r} c_{ig} d \log p_{ig}/p_{ir} \right), \\ &= \tilde{\lambda}_i \left(\sum_{g \in \mathcal{G}} s_{ig} d \log p_{ig}/p_{ir} - \sum_{g \in \mathcal{G}} c_{ig} d \log p_{ig}/p_{ir} \right), \\ &= \tilde{\lambda}_i \left(\sum_{g \in \mathcal{G}} \left(s_{ig} - c_{ig} \right) d \log p_{ig} \right), \\ &= \tilde{\lambda}_i \left(\sum_{g \in \mathcal{G}} \left(s_{ig} - \frac{\frac{q_{ig} mc_{ig}}{C(y,p)}}{s_{ig}} s_{ig} \right) d \log p_{ig} \right), \\ &= \tilde{\lambda}_i \left(\sum_{g \in \mathcal{G}} \left(s_{ig} - \frac{1}{\gamma_{ig}} s_{ig} \right) d \log p_{ig} \right), \\ &= \tilde{\lambda}_i \left(E_{\boldsymbol{s}_i} \left[d \log \boldsymbol{p}_i \right] E_{\boldsymbol{s}_i} \left[\frac{1}{\boldsymbol{\gamma}_i} \right] - E_{\boldsymbol{s}_i} \left[d \log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i} \right] \right), \end{aligned}$$

$$= -\tilde{\lambda}_i Cov_{\boldsymbol{s}_i}\left(d\log \boldsymbol{p}_i, \frac{1}{\boldsymbol{\gamma}_i}\right).$$

where $c_{ig} = \frac{q_{ig} mc_{ig}}{C(\boldsymbol{q}_i, \boldsymbol{p}_i)}$.

□

Proof of Proposition 4

Proof. From resource constraint,

$$q_{ig} = y_{ig} + \sum_{j \in N} x_{jig}.$$

$$d\log y_{ig} = \frac{q_{ig}}{y_{ig}} d\log q_{ig} - \sum_j \frac{x_{jig}}{y_{ig}} d\log x_{jig}.$$

From cost minimization of joint production, we know

$$\sum_g q_{ig} mc_{ig} d\log q_{ig} = \sum_{jp} x_{i,jp} p_{jp} d\log x_{i,jp} + \sum_f w_f x_{i,f} d\log L_{if}.$$

$$\sum_{jp} \frac{x_{i,jp} p_{jp}}{GDP} d\log x_{i,jp} = \sum_g \frac{1}{\mu_{ig}} \frac{q_{ig} p_{ig}}{GDP} d\log q_{ig} - \sum_f \frac{w_f x_{i,f}}{GDP} d\log L_{if}.$$

By definition,

$$\begin{aligned}
d\log Y &= \sum_{ig} \frac{p_{ig} y_{ig}}{GDP} d\log y_{ig} - \sum_f \frac{w_f x_{i,f}}{q_{ir} mc_{ig}} d\log L_f, \\
&= \sum_{ig} \frac{p_{ig} y_{ig}}{GDP} \left(\frac{q_{ir}}{y_{ig}} d\log q_{ig} - \sum_j \frac{x_{jig}}{y_{ig}} d\log x_{jig} \right) - \sum_f \frac{w_f L_f}{GDP} d\log L_f, \\
&= \sum_{ig} \frac{p_{ig} y_{ig}}{GDP} \frac{q_{ir}}{y_{ig}} d\log q_{ig} - \sum_{ig} \sum_j \frac{p_{ig} y_{ig}}{GDP} \frac{x_{jig}}{y_{ig}} d\log x_{jig} - \sum_f \frac{w_f L_f}{GDP} d\log L_f, \\
&= \sum_{ig} \frac{p_{ig} q_{ig}}{GDP} d\log q_{ig} - \sum_{ig} \frac{1}{\mu_{ig}} \frac{p_{ig} q_{ig}}{GDP} d\log q_{ig} + \sum_{if} \frac{w_f x_{i,f}}{GDP} d\log L_{if} - \sum_f \frac{w_f L_f}{GDP} d\log L_f, \\
&= \sum_{ig} \lambda_{ig} \left(1 - \mu_{ig}^{-1}\right) d\log q_{ig}.
\end{aligned}$$

Therefore,

$$\frac{\partial \log Y}{\partial \log \mu_{ig'}} = \sum_{ig} \lambda_{ig} \left(1 - \mu_{ig}^{-1}\right) \frac{d \log q_{ig}}{d \log \mu_{jg'}},$$
$$\frac{\partial \log Y}{\partial \log \mu_{ig} \partial \log \mu_{jg'}} = \sum_{ig,jg'} \lambda_{ig} d \log \mu_{ig} \frac{d \log q_{ig}}{d \log \mu_{jg'}}.$$

$$\begin{aligned}\frac{\partial \log Y}{\partial \log \mu_{ig} \partial \log \mu_{jg'}} &= \sum_{ig} \sum_{jg'} \lambda_{ig} d \log \mu_{ig} d \log \mu_{jg'} \frac{d \log q_{ig}}{d \log \mu_{jg'}}, \\ &= \sum_{ig} \lambda_{ig} d \log q_{ig} d \log \mu_{ig}.\end{aligned}$$

□

Proof of Proposition 5

Proof. Define the firm-level Domar weight:

$$\lambda_i = \sum_g \lambda_{ig}.$$

Let $d \log \mu_{ig}$ denote the exogenous change in the markup for product g of firm i. The aggregate markup change for firm i is:

$$d \log \mu_i = \frac{1}{\lambda_i} \sum_g \lambda_{ig} \, d \log \mu_{ig}.$$

The weighted average markup change is:

$$d \log \bar{\mu} = \sum_i \lambda_i \, d \log \mu_i = \sum_i \sum_g \lambda_{ig} \, d \log \mu_{ig}.$$

Let $d \log p_{ig}$ be the change in the price of product g by firm i, and $d \log w$ be the change in the wage rate.

The aggregate price index change is:

$$d \log p = \sum_i \sum_g \lambda_{ig} \, d \log p_{ig}.$$

The price change for product g of firm i is:

$$d\log p_{ig} = d\log\mu_i + d\log w + \frac{\kappa}{\lambda_i}\left(d\log\mu_{ig} - d\log\mu_i\right),$$

where

$$\kappa = \frac{\sigma}{\sigma+1}.$$

Using the zero-profit condition and labor market clearing, we have:

$$d\log w = -d\,\bar{\log}\,\mu.$$

Therefore, the aggregate price index change simplifies to:

$$d\log p = \sum_i \lambda_i\, d\log\mu_i - d\,\bar{\log}\,\mu = 0,$$

so $d\log p = 0$.

The change in the Domar weight of product g is:

$$d\log\lambda_{ig} = (1-\theta)\, d\log p_{ig}.$$

The welfare change is:

$$\mathcal{L} = \frac{1}{2}\sum_i\sum_g \lambda_{ig}\, d\log q_{ig}\, d\log\mu_{ig} = -\frac{1}{2}\theta\sum_i\sum_g \lambda_{ig}\, d\log p_{ig}\, d\log\mu_{ig}.$$

Our goal is to express $\mathcal{L}$ in terms of $d\log\mu_{ig}$. Using $d\log w = -d\,\bar{\log}\,\mu$, we have:

$$d\log p_{ig} = (d\log\mu_i - d\,\bar{\log}\,\mu) + \frac{\kappa}{\lambda_i}\left(d\log\mu_{ig} - d\log\mu_i\right).$$

Let

$$d\,\tilde{\log}\,\mu_i = d\log\mu_i - d\,\bar{\log}\,\mu,$$

and

$$d\,\tilde{\log}\,\mu_{ig} = d\log\mu_{ig} - d\log\mu_i,$$

then:

$$d\log p_{ig} = d\,\tilde{\log}\,\mu_i + \frac{\kappa}{\lambda_i} d\,\tilde{\log}\,\mu_{ig}.$$

Substituting into the welfare change:

$$\mathcal{L} = -\frac{1}{2}\theta \sum_i \sum_g \lambda_{ig} \left(d\,\tilde{\log}\,\mu_i + \frac{\kappa}{\lambda_i} d\,\tilde{\log}\,\mu_{ig} \right) d\log\mu_{ig}.$$

This simplifies to:

$$\mathcal{L} = -\frac{1}{2}\theta \left(\sum_i \lambda_i d\,\tilde{\log}\,\mu_i d\log\mu_i + \kappa \sum_i \frac{1}{\lambda_i} \sum_g \lambda_{ig} d\,\tilde{\log}\,\mu_{ig}^2 \right).$$

By definitions of total variance and conditional variance, the welfare change becomes:

$$\mathcal{L} = -\frac{1}{2}\theta \left(\mathrm{Var}_{\bar{\lambda}}(d\log\mu_i) + \frac{\sigma}{\sigma+1} \sum_i \lambda_i \,\mathrm{Var}_{s_i}(d\log\mu_{ig}) \right).$$

By the law of total variance, we know

$$\mathrm{Var}_\Lambda(d\log\mu_{ig}) = \mathrm{Var}_\Lambda(d\log\mu_i) + \sum_i \lambda_i \,\mathrm{Var}_{s_i}(d\log\mu_{ig}).$$

Using this, we obtain

$$\mathcal{L} = -\frac{1}{2}\theta \left(\mathrm{Var}_\Lambda(d\log\mu_{ig}) - \frac{1}{\sigma+1} \sum_i \lambda_i \,\mathrm{Var}_{s_i}(d\log\mu_{ig}) \right).$$

$\square$