Review of Economic Studies (2026) **00**, 1–34 doi:10.1093/restud/rdag123
© The Author(s) 2026. Published by Oxford University Press on behalf of The Review of Economic Studies Limited. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
Advance access publication 11 November 2026

Competition in Markups: A Theory of the Lerner Index in Networks and General Equilibrium

Lingxuan Wu
NYU Stern, USA

First version received December 2022*; Accepted September* 2026 *(Eds.)*

What determines demand elasticities and markups in production networks, where firms' demand and pricing are interdependent? I propose *competition in markups*, an equilibrium concept in which firms choose markups, taking other firms' markups as given. In partial equilibrium with vertical structures, it links upstream demand elasticities to downstream characteristics without backward induction and ties pass-through to the curvature of final demand. In general equilibrium with arbitrary industry networks, it microfounds firms' demand elasticities and markups through their within-industry market shares and four industry-level network statistics—cycle, substitution in production, substitution in consumption, and consumer-price impact. Empirical analysis of Compustat firms from 1997 to 2019 confirms the predicted variation of industry demand elasticities and firm markups. Relative to an exogenous-markup benchmark, endogenous markups dampen the aggregate total factor productivity impact of a productivity shock to an industry's largest firm by 6% for the median industry, and by up to 48% across industries.

Key words: Lerner index, General equilibrium, Production networks, Endogenous markups, Sectoral heterogeneity, Aggregate productivity

JEL codes: C72, D2, D43, D5, E1, L16, O47

1. INTRODUCTION

Understanding market power has been central to economics since its inception. In seminal work, Lerner (1934) proposes an index of market power based on markups and links it to the inverse of the demand elasticity. In simple cases where firms sell to consumers, it is common practice to postulate a demand function dependent on prices. In production networks, however, even if we specify a final demand function for the most downstream firms, what determines the demand elasticities for upstream firms? This challenge is compounded in general equilibrium models, where the primitives are production functions and preferences, not explicit demand functions.

The editor in charge of this paper was Thomas Chaney.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Outside special cases such as monopolistic competition or models with a continuum of industries (Atkeson and Burstein, 2008), how do we microfound demand elasticities and markups?

This paper proposes a new modelling approach of *competition in markups*, which handles arbitrary network structures and is suitable for both partial and general equilibrium analysis. This approach postulates that firms simultaneously choose markups, meaning that each firm assumes other firms hold their markups fixed when setting its own markup. In partial equilibrium (PE) with given final demand functions, competition in markups is equivalent to competition in prices in a horizontal economy with constant marginal costs and is more plausible and tractable in the presence of vertical structures. In general equilibrium (GE), competition in markups microfounds demand elasticities for all firms analytically in terms of their network positions. The theory predicts, and the data confirm, that firms' markups depend on their market shares and four industry-level network statistics: cycle, substitution in production, substitution in consumption, and consumer-price impact. This approach further allows us to quantify how endogenous markups affect shock propagation in the economy. Relative to an exogenous-markup benchmark, endogenous markups dampen the aggregate total factor productivity (TFP) impact of a productivity shock to an industry's largest firm in Compustat by 6% for the median industry, and by up to 48% across industries.

I develop the theory of competition in markups in two steps. I first illustrate its properties with given final demand functions in PE and then show how it microfounds demand elasticities in GE. In each setting, I highlight the limitations and difficulties of the canonical approach of competition in prices and explain how they are resolved by competition in markups. In Section 2, I use three examples in PE to show that competition in markups is a distinct equilibrium concept that coincides with competition in prices in horizontal economies and extends tractably to vertical structures with more desirable properties. To start with, in a horizontal economy where each firm has a constant marginal cost, I clarify that competition in markups is equivalent to competition in prices. That is because when the competitors' marginal costs are independent of one's action, assuming that they hold fixed their markups amounts to assuming that they hold fixed prices.

I then show that competition in markups delivers more realistic and tractable predictions than competition in prices when vertical structures are involved. As a second example, I consider a tree economy, where a group of upstream firms sell to one downstream firm that then sells to the final customer. Competition in markups predicts that upstream markups depend on downstream characteristics, which is empirically realistic, whereas competition in prices predicts otherwise. Under competition in markups, in setting its markup, an upstream firm anticipates two margins—the *conditional input demand* and the *final demand*—that jointly determine its demand elasticity. First, raising its markup makes its product more expensive relative to other upstream firms, which triggers the downstream firm to adjust its *conditional input demand* across upstream products. Second, raising its markup increases the downstream marginal cost and thus the downstream price, since the downstream firm is assumed to hold fixed its markup, resulting in a decline in the *final demand*. Consequently, through the final demand margin, shocks to the downstream firm's productivity or the final demand generally affect upstream markups. In contrast, competition in prices shuts down the final demand margin by assuming that the downstream firm holds fixed its price. As a result, downstream productivity or demand shocks do *not* affect upstream markups.

Crucially, competition in markups assumes a perceived pass-through of one—each firm expects its price changes to pass fully into downstream prices—but does not imply an equilibrium pass-through of one. Analogously, competition in prices assumes a perceived pass-through of zero, yet its equilibrium pass-through is generally non-zero. To characterize the equilibrium pass-through, I construct a third example of an N-firm vertical production chain. Competition

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

in markups links all markups to the final demand elasticity and pass-through to its curvature: a productivity shock to any firm passes into the final price incompletely if and only if the final demand elasticity increases with the final price. This is the same condition as under monopoly pricing, regardless of the chain's length. Competition in prices compares unfavourably: its simultaneous-move version admits no equilibrium in a chain,[1] while its sequential-move version—an N-stage backward induction in which each firm anticipates the entire chain of downstream responses—predicts markups and pass-through that depend on up to the $(N+1)$th-order elasticity of the final demand, an object with little empirical counterpart.

After illustrating the properties of competition in markups in PE, I show in Section 3 that it seamlessly integrates with GE and microfounds demand elasticities and markups. Once we move from PE to GE, there is no longer a given final demand function, and competition in prices is not a straightforward solution concept as prices arise endogenously from market clearing. I integrate competition in markups into GE analysis using a two-stage game. In the first stage, all firms simultaneously choose markups to maximize their profits, which are determined in the second stage. In the second stage, all firms minimize input costs given the markup profile chosen in the first stage, consumers optimize, and markets clear, allowing prices and quantities to emerge endogenously. The second stage is a standard GE setup that treats markups as exogenous wedges. It gives rise to profit functions that depend on all firms' markups. These profit functions microfound the demand elasticities in GE based on network structures and in turn generalize the Lerner rule to GE by linking markups inversely to these elasticities.

To deliver concrete predictions on GE demand elasticities, I consider a nested industry-firm constant-elasticity-of-substitution (CES) economy with arbitrary industry networks.[2] Within each industry, firm products aggregate into CES industry composites with elasticity ρ. Firms in the same industry use the same CES production technologies combining labour and all industry composites with elasticity θ_p and differ in productivities. A representative consumer with quasi-linear utility supplies labour and consumes a CES bundle of industry composites with elasticity θ_c, where the curvature of consumption utility is parametrized by ϵ.

I analytically characterize the demand elasticities and markups of all firms in this network. I show that the demand elasticity of firm i_n in industry n is the market-share-weighted average of an "industry demand elasticity" η_n and the within-industry elasticity of substitution ρ, plus an adjustment term related to its cost-based market share in the industry. This generalizes the formula in Atkeson and Burstein (2008) by endogenously generating a variable industry demand elasticity η_n, which captures the sensitivity of demand for the industry composite to the industry markup. I establish that the demand elasticity η_n varies across industries, depending additively on four network statistics:

$$\eta_n = \text{cycle}_n + \left(\theta_p - 1\right) \text{SP}_n + \left(\theta_c - 1\right) \text{SC}_n + \left(\epsilon - 1\right) \text{CP}_n + 1. \tag{1}$$

Here, θ_p and θ_c are the elasticities of substitution across industry composites in production and in consumption, and ϵ is the curvature parameter of consumption utility. The four network statistics depend only on the input–output structure. The *cycle* (cycle_n) measures how much of industry n's own output feeds back into its cost through the supply chain. *Substitution in production* (SP_n) and *substitution in consumption* (SC_n) measure how unevenly industry n's product is used

1. With a single buyer whose price is held fixed, the conditional-input-demand margin is absent and the final-demand margin is shut by assumption, so perceived demand is perfectly inelastic.

2. I show in Supplementary Appendix D that competition in markups applies to arbitrary firm networks with arbitrary constant-returns production functions and that the predictions extend with the four network statistics generalized accordingly.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

across the inputs of its downstream firms and across the consumer's basket, which governs how strongly their expenditure shifts towards (away from) it if $\theta_p, \theta_c < 1$ (> 1). The *consumer-price impact* (CP_n) is the elasticity of the consumer price index with respect to industry n's markup. With quasi-linear utility, aggregate consumption falls by $(\epsilon - 1)$ per cent when the consumer price index rises by 1%, which scales down every firm's demand proportionally.

The theory implies that controlling for the market share, firms in industries that are less elastic set higher markups. Furthermore, under the realistic parameter restriction $\eta_n < \rho$, larger firms in each industry set higher markups, and the slope of markup against market share varies across industries driven by η_n. I test this prediction using Compustat data in Section 5. It is recognized that markups estimated from Compustat revenue data are not the true markups, but their log variations within an industry are informative about the true log variations (De Ridder *et al.*, 2026). Thus I consider a log–log approximation of the markup-on-market-share relation, and show that such a slope varies systematically across industries based on the four network statistics as suggested by (1).

As an application, in Section 6, I revisit the aggregate TFP response to firm-level productivity shocks. I show that relative to an exogenous-markup benchmark, endogenous markups dampen the aggregate TFP impact of a productivity shock to an industry's largest firm. As the top firm gets more productive, its market share grows and it charges a higher markup. This leads to a larger increase in industry markup and a smaller increase in industry productivity relative to the exogenous-markup benchmark, hence dampening the TFP response. Across industries in 2019, the aggregate TFP response is dampened by 6% for the median industry, and by up to 48% across industries.

The paper is organized as follows. Section 2 illustrates the properties of competition in markups in PE. Section 3 introduces competition in markups to GE. Section 4 derives the variable industry demand elasticities in a nested CES setup, and Supplementary Appendix D shows the applicability of our approach to arbitrary firm network structures and constant-returns production functions. Section 5 tests the theoretical prediction on firms' markups driven by variable industry demand elasticities. Section 6 quantifies the aggregate TFP response to firm-level productivity shocks. Section 7 concludes with future directions. All appendices, including Supplementary Appendix F that collects all proofs, are in the Supplementary Appendix.

1.1. *Literature review*

This paper contributes to the literature on market power in production networks with a model that microfounds demand elasticities and markups in a general network. Atkeson and Burstein (2008) consider a horizontal economy (*i.e.* with no intermediate inputs) with a continuum of industries and predict that larger firms charge higher markups. Because industries are infinitesimal, the industry demand elasticity is constant and firms' markups only depend on their market shares. I consider a general production network with a finite number of firms and industries à la Baqaee and Farhi (2020). My theory predicts—and the data confirm—that firms' markups depend not only on their market shares but also on an industry demand elasticity that varies systematically across industries.[3] Grassi (2017) and Dhyne *et al.* (2022) also consider production networks. However, their models assume that when upstream firms make their pricing decisions, they do

3. Cross-industry studies in the field of industrial organization that involve regressing markups or profits on concentration trace back to Bain (1951, 1956), known as the "structure–conduct–performance" paradigm. The paradigm was later discredited as researchers realized that the variables involved are all endogenous, so that such associations cannot be interpreted without a model (Bresnahan, 1989; Schmalensee, 1989). This paper provides such a model and distinguishes "structural" relations, which can be tested using observed market shares and industries' network positions

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

not internalize the responses of downstream prices or quantities. As a result, markups in their frameworks depend solely on market shares, which are determined only by productivities and CES aggregation *within* an industry. It follows that any other shock (such as productivity shocks to other industries or demand shocks) leaves markups unchanged. In contrast, my model allows markups of all firms to respond to all shocks.[4]

Efforts to understand market power in GE go back decades. To the best of my knowledge, however, this research typically entails restrictions on the input–output network structures and production functions. Some scholarship, such as Gabszewicz and Vial (1972) and FitzRoy (1974), and recently Azar and Vives (2021), studies GE with competition in quantities. Other work, including Marschak and Selten (1974), Nikaido (1975), and Benassy (1988), studies competition in prices. My model of competition in markups shares similarities with competition in prices and handles arbitrary network structures and production functions. It seamlessly integrates with GE as it allows prices and quantities to arise endogenously. Pellegrino (2025) and Ederer and Pellegrino (2026) consider imperfect competition under a rich hedonic demand system of a representative consumer. This paper complements theirs by allowing for a rich input–output structure and microfounding intermediate demand functions in GE. Finally, my theory is an explicit model of a non-cooperative game that firms play and explains markups in the data. This contrasts with the subjective approach (Negishi, 1961, 1972), which can lack predictive power (Hart, 1985).[5]

Work on production networks dates back to Leontief (1936) and Long and Plosser (1983). More recently, Gabaix (2011), Acemoglu *et al.* (2012), and Baqaee and Farhi (2019) study the propagation of microeconomic shocks in efficient networks. Jones (2011, 2013), Liu (2019), Baqaee and Farhi (2020), and Bigio and La'O (2020) study efficiency implications of *exogenous* wedges in networks. In contrast, I endogenize markups and study how endogenous markups influence the aggregate TFP responses to firm-level productivity shocks. Acemoglu and Tahbaz-Salehi (2025), Voronina (2022), Alviarez *et al.* (2026), and Bizzarri (2026) study two-sided market power via bilateral bargaining or competition in supply and demand functions, under certain restrictions on the network structure or production function. My model applies to an arbitrary network structure and delivers testable implications for markups based on firms' network positions.

My model makes testable predictions on pass-through, which a large empirical literature measures. Studies have found that cost shocks pass through into prices incompletely (Feenstra *et al.*, 1996; Goldberg and Hellerstein, 2013; Amiti *et al.*, 2019; Ganapati *et al.*, 2020; Auer *et al.*, 2021; Cavallo *et al.*, 2021). Combined with the evidence that demand elasticities increase with price (*e.g.* Nakamura and Zerom, 2010), these findings are consistent with my model's prediction in the partial-equilibrium vertical economy: cost shocks pass into the final price incompletely precisely when the final demand elasticity increases with the final price. It may be fruitful for future work to test the model's pass-through predictions across the production network more broadly.

(Propositions 5 and 6), from relations that suffer from omitted-variable bias (Proposition E1 in the Supplementary Appendix).

4. A further conceptual difference is that their frameworks require a nested industry-firm structure, whereas my solution concept applies to arbitrary network structures and constant-returns production functions (as shown in Supplementary Appendix D) without the need to define industries.

5. PE models with conjectural variations that nest competition in quantities or prices date back to Bowley (1924). Salinger (1989) studies vertical relations with conjectural variations.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

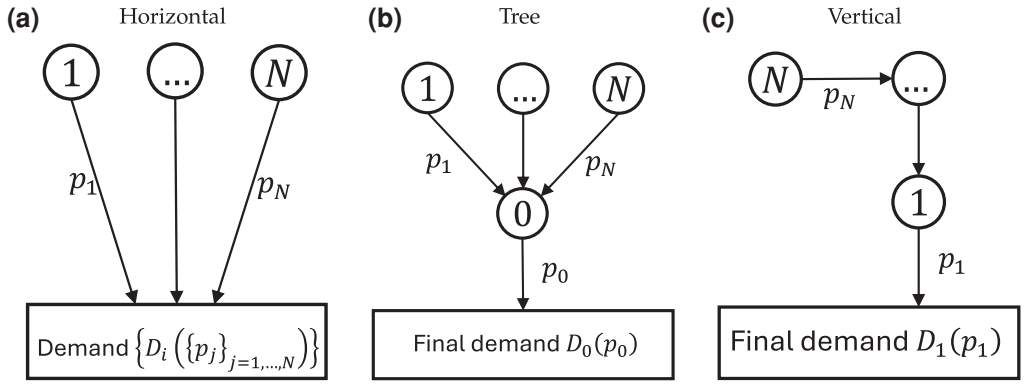


FIGURE 1
Examples in partial equilibrium: (a) horizontal economy; (b) tree economy; (c) vertical economy

2. PARTIAL EQUILIBRIUM WITH COMPETITION IN MARKUPS

In this section, I illustrate the properties of competition in markups (*i.e.* simultaneous moves in markups) under a given final demand function. I compare it with two traditional solution concepts: simultaneous price-setting and sequential price-setting (where upstream firms move before downstream firms). The latter is a common assumption in models with a vertical structure (Spengler, 1950; Tirole, 1988, Section 4.2.2). For brevity, I refer to competition in markups as SIM-μ, simultaneous price-setting as SIM-p, and sequential price-setting as SEQ-p.

I consider three examples, illustrated in Figure 1: a horizontal economy, a tree economy, and a vertical economy. In the horizontal economy, I demonstrate that SIM-μ is equivalent to SIM-p. In the tree economy, SIM-μ is more realistic than SIM-p, in that SIM-μ allows downstream demand and productivity shocks to affect upstream markups. In the vertical economy, where SIM-p becomes ill-defined and SEQ-p is the traditional solution, I show that SIM-μ is more tractable and delivers more intuitive properties of pass-through.

2.1. *Horizontal economy*

I start with a simple horizontal economy, depicted in Figure 1(a), and show that competition in markups (SIM-μ) is equivalent to simultaneous price-setting (SIM-p).

Specifically, there are N firms indexed by $i \in \{1, \ldots, N\}$, each with a constant-returns production function $y_i = z_i l_i$ with the wage normalized to one (and hence a constant marginal cost z_i^{-1}) and facing a demand function $D_i(\{p_j\}_{j\in\{1,\ldots,N\}})$. In this case, when firm i considers its pricing, assuming other firms hold their markups fixed amounts to assuming other firms hold their prices fixed, since p_i simply equals $\mu_i z_i^{-1}$. Hence, SIM-μ and SIM-p are equivalent.

Proposition 1 (PE horizontal economy)**.** *In the horizontal economy in PE (Figure 1(a)), competition in markups (SIM-μ) is equivalent to simultaneous price-setting (SIM-p).*

Specifically, under both assumptions, the markups charged by firm $i \in \{1, \ldots, N\}$ satisfy

$$\underbrace{\frac{\mu_i}{\mu_i - 1}}_{\text{inverse of Lerner index}} = \underbrace{-\frac{\partial \log D_i\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right)}{\partial \log p_i}}_{\text{elasticity of demand function}}. \tag{2}$$

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Further, the responses of prices $\{p_i\}_{i\in\{1,\ldots,N\}}$ *and markups* $\{\mu_i\}_{i\in\{1,\ldots,N\}}$ *to productivity shocks* $\{\mathrm{d}z_i\}_{i\in\{1,\ldots,N\}}$ *are the same across SIM-*μ *and SIM-p.*

2.2. *Tree economy*

Now I introduce a vertical relation to form a tree economy (Figure 1(b)), and show that competition in markups (SIM-μ) makes more realistic predictions than simultaneous price-setting (SIM-p): The final demand elasticity and downstream productivity shocks affect upstream markups under SIM-μ but not SIM-p.

Similar to the horizontal economy, there are N upstream firms indexed by $i \in \{1, \ldots, N\}$, each with a constant-returns production function $y_i = z_i l_i$ with the wage normalized to one (and hence a constant marginal cost z_i^{-1}). In addition, there is one downstream firm indexed by 0 with a constant-returns production function $z_0 F_0(\{x_i\}_{i\in\{1,\ldots,N\}})$.[6] Downstream firm 0 faces a final demand function $D_0(p_0; \zeta)$ that is only dependent on its price, subject to a demand shock ζ.

Downstream firm 0 minimizes its cost

$$\min \sum_{i=1}^{N} p_i x_i, \text{ s.t. } y_0 = z_0 F_0\left(\{x_i\}_{i\in\{1,\ldots,N\}}\right).$$

It gives rise to a constant marginal cost $z_0^{-1}\mathbf{C}_0(\{p_j\}_{j\in\{1,\ldots,N\}})$ and conditional input demand functions $\{D_i(\{p_j\}_{j\in\{1,\ldots,N\}}, z_0^{-1}y_0)\}_{i\in\{1,\ldots,N\}}$ for upstream firms that are linear in downstream output y_0, which equals $D_0(p_0; \zeta)$. Hence the demand for upstream firm i is

$$D_i = \underbrace{d_i\left(\{p_j\}_{j\in\{1,\ldots,N\}}\right)}_{\text{unit input demand}} \cdot \underbrace{z_0^{-1}\overbrace{D_0(p_0;\zeta)}^{\text{final demand}}}_{\text{quantity of input bundle}} \tag{3}$$

where $d_i(\{p_j\}_{j\in\{1,\ldots,N\}}) = \frac{\partial \mathbf{C}_0(\{p_j\}_{j\in\{1,\ldots,N\}})}{\partial p_i}$ is the conditional input demand per unit of input bundle due to Shephard's lemma. $d_i(\cdot)$ is independent of the downstream firm's productivity z_0 or its desired output y_0 since the production function is constant-returns.

Proposition 2 (PE tree economy)**.** *Consider the tree economy in PE (Figure 1(b)).*

(1) *Under competition in markups (SIM-*μ*), downstream firm 0 charges a markup of*

$$\underbrace{\frac{\mu_0}{\mu_0 - 1}}_{\text{inverse of Lerner index}} = \underbrace{-\frac{\mathrm{d}\log D_0(p_0;\zeta)}{\mathrm{d}\log p_0}}_{\text{final demand elasticity}} \tag{4}$$

and upstream firm $i \in \{1, \ldots, N\}$ *charges a markup of*

$$\frac{\mu_i}{\mu_i - 1} = \underbrace{-\frac{\partial \log d_i\left(\{p_j\}_{j\in\{1,\ldots,N\}}\right)}{\partial \log p_i}}_{\text{conditional input demand elasticity}} + \underbrace{\left(-\frac{\mathrm{d}\log D_0(p_0;\zeta)}{\mathrm{d}\log p_0}\right)}_{\text{final demand elasticity}} \cdot \underbrace{\frac{p_i d_i}{\sum_{j\in\{1,\ldots,N\}} p_j d_j}}_{\text{firm } i\text{'s share in firm 0's cost}} \tag{5}$$

6. Our prediction that upstream markups depend on downstream characteristics under competition in markups but not competition in prices holds as long as the downstream production function is homothetic.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

which depends on the final demand elasticity. Consequently, downstream productivity shock z_0 and demand shock ζ generally affect both downstream and upstream markups.

(2) *Under simultaneous price-setting (SIM-p), downstream firm 0 charges a markup that satisfies (4) and upstream firm $i \in \{1, \ldots, N\}$ charges a markup of*

$$\frac{\mu_i}{\mu_i - 1} = \underbrace{-\frac{\partial \log d_i\left(\{p_j\}_{j\in\{1,\ldots,N\}}\right)}{\partial \log p_i}}_{\text{conditional input demand elasticity}} \tag{6}$$

which is independent of the final demand elasticity. Consequently, downstream productivity shock z_0 and demand shock ζ only affect the downstream markup.

Under both assumptions, an upstream firm i maximizes its profits $\pi_i = (\mu_i - 1)z_i^{-1}D_i$ by setting a markup satisfying

$$\underbrace{\frac{\mu_i}{\mu_i - 1}}_{\text{inverse of Lerner index}} = \underbrace{-\frac{\mathrm{d}\log D_i\left(\{p_j\}_{j\in\{1,\ldots,N\}}, z_0^{-1}y_0\right)}{\mathrm{d}\log p_i}}_{\text{demand elasticity}}.$$

However, the demand elasticity differs, as an upstream firm anticipates that changing its price p_i affects final demand y_0 under competition in markups, but not under competition in prices.

Under competition in markups (SIM-μ), when each upstream firm i considers its markup μ_i, it assumes that the downstream firm 0 holds its markup fixed and thus increases its price p_0 if μ_i goes up. As a result, the final demand $y_0 = D_0(p_0; \zeta)$ will decline, which lowers demand for firm i's product via the final demand term in (3). Its magnitude is governed by the price sensitivity of final price p_0 to upstream price p_i, which equals the cost share of input i in firm 0's production according to Shephard's lemma. Further, as firm i's product becomes more expensive, the downstream firm reoptimizes its input demands across upstream firms, *i.e.* a change via the unit input demand $d_i(\cdot)$ in (3). These two margins together constitute the elasticity of firm i's demand $\left(-\frac{\partial \log D_i}{\partial \log p_i}\right)$, hence giving rise to the pricing formula (5). Moreover, if the downstream firm 0 experiences a productivity shock, its price changes and the final demand elasticity $\left(-\frac{\mathrm{d}\log D_0(p_0;\zeta)}{\mathrm{d}\log p_0}\right)$ generally changes (unless D_0 entails a constant elasticity). As a result, upstream markups are affected by downstream productivity shocks. By the same logic, a shock to the final demand ζ also affects the upstream markups through the final demand term in (3). It is in line with findings in Sangani (2025) that upstream firms charge higher markups when supplying to retailers facing more inelastic final demand.

In contrast, under simultaneous price-setting (SIM-p), an upstream firm i does not internalize its impact on downstream price p_0 nor downstream demand y_0. Hence, the elasticity of firm i's demand only arises from the input choices made by downstream firm 0, *i.e.* the unit input demand $d_i(\cdot)$ in (3). The optimal markup is characterized by the pricing formula (6), which only features the elasticity of the unit input demand $\left(-\frac{\partial \log d_i}{\partial \log p_i}\right)$ but not the final demand elasticity $\left(-\frac{\mathrm{d}\log D_0}{\mathrm{d}\log p_0}\right)$. Consequently, downstream productivity and demand shocks leave upstream markups unchanged.

The relevance of downstream demand and productivity shocks for upstream markups under SIM-μ stems from the assumption that upstream firms perceive a downstream pass-through of one. Crucially, it does not imply that the equilibrium pass-through of upstream cost shocks to

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

downstream price is one—a point we elaborate further in the next subsection. The idea is important yet simple: under a Nash equilibrium, each player assumes that opponents hold their actions fixed when choosing their own action, given the payoff function. However, when the payoff functions change (for example, due to productivity shocks), the equilibrium actions generally change as well.

2.2.1. Invariance of SIM-μ to inserting competitive intermediaries. Finally, we note that the SIM-μ pricing formula is invariant to inserting zero-markup competitive intermediaries between the N oligopolists and the final customer, whereas SIM-p is not. To see this, suppose the downstream consists of a continuum of undifferentiated retailers, who therefore price at marginal cost, $p_0 = z_0^{-1}\mathbf{C}_0(\{p_j\}_{j\in\{1,\dots,N\}})$. This tree economy is then economically identical to the horizontal economy of Section 2.1, with competitive retailers that merely pass costs through. So it is desirable that an equilibrium concept predict the same markups in both economies; otherwise, predictions would hinge on the arbitrary modelling choice of whether to include the competitive retailers explicitly. The optimal upstream markups (5) under SIM-μ pass this test: substituting demand $D_i(\{p_j\}_{j\in\{1,\dots,N\}}) = d_i(\{p_j\}_{j\in\{1,\dots,N\}}) \cdot z_0^{-1} D_0(z_0^{-1}\mathbf{C}_0(\{p_j\}_{j\in\{1,\dots,N\}}); \zeta)$ in Proposition 1 yields exactly (5) in Proposition 2. It happens because the unit pass-through that SIM-μ perceives is exactly the competitive retailers' actual pass-through. The upstream markups under SIM-p (6) do not, because holding the retailers' prices fixed severs the final-demand margin.

2.3. *Vertical economy*

Last but not least, we consider a vertical economy (Figure 1(c)). The traditional solution concept in this setting is sequential price-setting (SEQ-p) where upstream firms move before downstream firms (Spengler, 1950; Tirole, 1988, Section 4.2.2). I show that competition in markups (SIM-μ) offers better tractability and avoids dependence of markups and pass-through on higher-order elasticities of final demand, which are challenging to measure.

Specifically, there are N firms indexed by $i \in \{1, \dots, N\}$ along a vertical chain. The most upstream firm N operates a constant-returns production function $y_N = z_N l_N$ with the wage normalized to one, whereas each downstream firm $i \in \{1, \dots, N-1\}$ operates $y_i = z_i x_i$ and sources from its upstream (*i.e.* $x_i = y_{i+1}$). A cost shock is a negative shock to productivity, *i.e.* $\mathrm{d}\log z_i^{-1}$. The most downstream firm 1 faces a final demand $D_1(p_1)$ that depends only on its price.

Simultaneous price-setting becomes ill-defined in this vertical economy. Since each downstream firm i only sources from its immediate upstream firm $i+1$, the conditional input demand for firm $(i+1)$'s product only depends on the quantity firm i seeks to produce, not on the upstream price p_{i+1}. If the upstream firm $i+1$ takes the downstream price p_i as given, it perceives an inelastic conditional input demand and thus charges an infinite markup. This pathological outcome has motivated researchers to adopt alternative assumptions, such as sequential price-setting.

Under sequential price-setting, firm N sets its price first, followed by firm $N-1$, and so on, constituting an N-stage dynamic game. As each firm contemplates its own price, it internalizes that raising its price increases the marginal cost of the downstream firm, which then raises the downstream price, ultimately reducing demand. As a result, each firm faces a downward-sloping demand curve and chooses a finite markup.

The following proposition shows that in this vertical economy, competition in markups also gives rise to a downward-sloping demand curve, by assuming that each firm perceives a pass-through of one. In addition, it offers better tractability and avoids dependence on higher-order elasticities of final demand.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Proposition 3 (PE vertical economy)**.** *Consider the vertical economy in PE (Figure 1(c)).*

(1) *Under competition in markups (SIM-μ), firm $i \in \{1, \ldots, N\}$ charges*

$$\underbrace{\frac{\mu_i}{\mu_i - 1}}_{\text{inverse of Lerner index}} = e^{(1)} \tag{7}$$

in which $e^{(1)} = -\frac{\mathrm{d}\log D_1}{\mathrm{d}\log p_1}$ is the final demand elasticity. In response to productivity shocks, the final price response depends only on the second-order elasticity of final demand as

$$\mathrm{d}\log p_1 = \frac{\sum_{i=1}^{N} \mathrm{d}\log z_i^{-1}}{1 + \frac{N}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}} \tag{8}$$

and features positive and incomplete pass-through, i.e. $\frac{\mathrm{d}\log p_1}{\mathrm{d}\log z_i^{-1}} \in (0, 1)$ for any i, if and only if $e^{(1)}$ increases with price p_1.

(2) *Under sequential price-setting (SEQ-p) where upstream firms move earlier, firm $i \in \{1, \ldots, N\}$ charges*

$$\frac{\mu_i}{\mu_i - 1} = e^{(i)} \tag{9}$$

in which $e^{(i)}$ involves the ith-order elasticity of the final demand $D_1(p_1)$, defined iteratively as follows, starting with $e^{(1)} = -\frac{\mathrm{d}\log D_1}{\mathrm{d}\log p_1}$,

$$e^{(i)} = \frac{e^{(1)}}{1 + \sum_{j=1}^{i-1} \frac{1}{e^{(j)}-1}\frac{\mathrm{d}\log e^{(j)}}{\mathrm{d}\log p_1}}$$

In response to productivity shocks, the final price response depends on up to $(N+1)$th-order elasticity of final demand as

$$\mathrm{d}\log p_1 = \frac{\sum_{i=1}^{N} \mathrm{d}\log z_i^{-1}}{1 + \sum_{j=1}^{N} \frac{1}{e^{(j)}-1}\frac{\mathrm{d}\log e^{(j)}}{\mathrm{d}\log p_1}}. \tag{10}$$

To understand these results, we start with the special case of monopoly pricing ($N = 1$), in which case SIM-μ and SEQ-p coincide. The markup charged by a monopoly follows

$$\frac{\mu_1}{\mu_1 - 1} = e^{(1)}$$

with $e^{(1)}$ being the final demand elasticity. Further, by taking total differentiation of this pricing formula, we have the pass-through of an increase in input cost as

$$\frac{\mathrm{d}\log p_1}{\mathrm{d}\log z_1^{-1}} = \frac{1}{1 + \frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}} \tag{11}$$

which is less than one if and only if the demand elasticity $e^{(1)}$ increases with price p_1 (Bulow and Pfleiderer, 1983; Mrázová and Neary, 2017).

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

With N firms, SIM-μ predicts that markups depend on the final demand elasticity and pass-through hinges on the final demand curvature. Each firm i anticipates that if it raises its price p_i, the final price p_1 will increase proportionally and the final demand will decline according to the final demand elasticity $e^{(1)}$, leading to the same proportional decline in i's demand. That is, each firm i behaves as if it directly faces the final demand elasticity. If the final demand elasticity increases with price, then firms' markup choices are strategic substitutes.

In terms of pass-through, an increase in any firm's marginal cost $\mathrm{d}\log z_i^{-1}$ would raise the final price p_1, which affects the final demand elasticity $e^{(1)}$ and, in turn, markups according to (7). Thus, if the final demand elasticity $e^{(1)}$ increases in price p_1, the pass-through of any productivity shock to final price will be incomplete as markups adjust to partially absorb the shock. This requirement for incomplete pass-through is the same as that in the standard monopoly problem (11) and the prediction aligns with findings that pass-through of upstream cost shocks to retail prices is incomplete (*e.g.* Cavallo *et al.*, 2021), and that demand elasticity increases with price (*e.g.* Nakamura and Zerom, 2010).

In contrast, with N firms, SEQ-p links markups and pass-through to higher-order elasticities of the final demand function. To understand these predictions, let us start with the case of two firms ($N = 2$), where firm 2 sets its price p_2, followed by firm 1. For any input price p_2, firm 1 is a monopoly facing the final demand, and thus its best response is given by the monopoly pass-through (11),

$$\frac{\mathrm{d}\log p_1}{\mathrm{d}\log p_2} = \frac{1}{1+\frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}}$$

which depends on the second-order elasticity of final demand. For firm 2, it internalizes firm 1's pass-through and thus perceives a demand elasticity of

$$e^{(2)} = -\frac{\mathrm{d}\log y_2}{\mathrm{d}\log p_2} = \underbrace{-\frac{\mathrm{d}\log D_1(p_1)}{\mathrm{d}\log p_1}}_{\text{final demand elasticity}} \cdot \underbrace{\frac{\mathrm{d}\log p_1}{\mathrm{d}\log p_2}}_{\text{downstream pass-through}} = \frac{e^{(1)}}{1+\frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}}.$$

Hence firm 2 sets a markup that depends on not only the final demand elasticity but also its second-order elasticity. The pass-through of any productivity shock into final price, when taking into account both firms' best responses, involves the derivative of firm 2's markup and hence incorporates the third-order demand elasticity.

With N firms, the most upstream firm anticipates $N-1$ best responses of downstream firms. Hence it faces a demand elasticity that depends on the Nth-order elasticity of final demand and sets its markup accordingly. Consequently, the pass-through of any productivity shock to final price involves up to $(N+1)$th-order demand elasticities.

In summary, both SIM-μ and SEQ-p allow upstream firms to anticipate downstream adjustments. SIM-μ achieves it with simultaneous moves, resulting in a system of N simultaneous equations, whereas SEQ-p solves an N-stage backward induction, which is more complex. Further, SIM-μ relates markups solely to final demand elasticity and pass-through to final demand curvature. Incomplete pass-through arises when the final demand elasticity increases with price, as in the standard monopoly problem. In contrast, SEQ-p determines markups and pass-through based on higher-order demand elasticities, which are difficult to measure.

Taking stock of this section, competition in markups is robust across various network structures given a demand function. In a horizontal economy, it is equivalent to simultaneous price-setting. When a vertical structure is present, it incorporates downstream adjustments without resorting to sequential moves, lets downstream productivity and demand shocks affect

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

upstream markups, and predicts intuitive properties of pass-through. Furthermore, as will become clear next, a general production network may involve cycles. In such cases, there is no natural way to specify a sequence of price-setting, yet competition in markups still applies.

3. GENERAL EQUILIBRIUM WITH COMPETITION IN MARKUPS

I set up a general network economy and introduce competition in markups into GE. Then I focus on a widely-studied nested CES network, which will allow us to separate within-industry substitution from variable industry demand elasticities.

3.1. *Setup and equilibrium definition*

The economy consists of a finite or infinite number of firms with arbitrary input–output linkages and a representative consumer that consumes and supplies labour. I use the wage as the numeraire.

The consumer consumes a constant-returns bundle of all goods in this economy indexed by $i \in \mathcal{T}$ (each produced by a firm)

$$C = C\left(\{c_i\}_{i\in\mathcal{T}}\right),$$

and maximizes utility $U(C, L)$, subject to the budget constraint

$$\sum_{i\in\mathcal{T}} p_i c_i = L + \sum_{i\in\mathcal{T}} \pi_i,$$

in which p_i is the price of good i, and π_i is a lump-sum rebate of profits from firm i. There exists a price index $P(\{p_i\}_{i\in\mathcal{T}})$ for the consumption bundle.

Each firm i uses a constant-returns production technology $y_i = z_i F_i(\{x_{ij}\}_{j\in\mathcal{T}}, l_i)$, which combines intermediate inputs $\{x_{ij}\}_{j\in\mathcal{T}}$ and labour l_i subject to productivity z_i, described by the cost function

$$z_i^{-1}\mathbf{C}_i\left(\left\{p_j\right\}_{j\in\mathcal{T}}\right) y_i$$

in which y_i is the total output and $\mathbf{C}_i$ is the unit cost of input bundle.

Definition 1. A *general equilibrium with competition in markups* is a pure-strategy Nash equilibrium among firms each with action μ_i to maximize its real profit

$$\pi_i^r\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right) = \underbrace{\left(1-\frac{1}{\mu_i}\right)}_{\text{profit margin}} \underbrace{R_i\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right)}_{\text{revenue}}, \quad \text{with}$$

$$R_i\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right) = \frac{p_i\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right) y_i\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right)}{P\left(\left\{\mu_j\right\}_{j\in\mathcal{T}}\right)} \tag{12}$$

where $p_i(\cdot)$ is the nominal price of good i, $y_i(\cdot)$ is the quantity sold of good i, and $P(\cdot)$ is the consumer price, all determined in GE as follows.

Given markups $\{\mu_i\}_{i\in\mathcal{T}}$, the GE is a system of prices and quantities $\{p_i, x_{ij}, y_i, l_i, c_i, L\}_{i,j\in\mathcal{T}}$ such that

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

(1) Each firm i chooses inputs $\{x_{ij}\}_{j\in\mathcal{T}}$ and l_i to minimize its cost and produces y_i, given markup μ_i on its marginal cost $z_i^{-1}\mathbf{C}_i$.
(2) The consumer chooses consumption $\{c_i\}_{i\in\mathcal{T}}$ and labour supply L to maximize utility, subject to the budget constraint where profits are rebated lump sum.
(3) The goods and labour markets clear, $c_i + \sum_{j\in\mathcal{T}} x_{ji} = y_i, \forall i \in \mathcal{T}$, and $\sum_{i\in\mathcal{T}} l_i = L$.

This is a two-stage definition, where I build competition in markups on top of GE. Competition in markups is a Nash equilibrium in markups given profit functions $\{\pi_i^r(\mu_i, \{\mu_j\}_{j\in\mathcal{T}\setminus\{i\}})\}_{i\in\mathcal{T}}$: each firm i's markup μ_i maximizes its profit π_i^r given all other firms' markups $\{\mu_j\}_{j\in\mathcal{T}\setminus\{i\}}$, so that the equilibrium is a fixed point of the firms' best responses in the space of markups. The profit functions derive from a second-stage GE in which firms take the markup profile $\{\mu_i\}_{i\in\mathcal{T}}$ as given and minimize costs, the consumer optimizes, and prices and quantities are pinned down through market clearing. All second-stage outcomes—prices, quantities, and hence profit functions $\{\pi_i^r\}_{i\in\mathcal{T}}$—are therefore functions of the markup profile $\{\mu_i\}_{i\in\mathcal{T}}$. Because firms' strategies are markups rather than prices or quantities, the definition is consistent with GE: prices and quantities are not chosen by any firm but emerge from market clearing. Accordingly, every partial derivative with respect to a markup in this paper refers to this second-stage mapping. In its first-order condition $0 = \frac{\partial \log \pi_i^r}{\partial \log \mu_i} = \frac{1}{\mu_i - 1} + \frac{\partial \log R_i}{\partial \log \mu_i}$, firm i varies its own markup, holds all other firms' markups fixed, and lets all prices and quantities adjust endogenously through the second-stage GE.

It should be immediately clear that the choice of numeraire in GE does not affect the equilibrium markups or quantities, since the profit function (12) is in real terms and thus invariant to the numeraire. The use of real profit is necessary because, in GE, a firm owner should not be concerned with nominal profit *per se*, but with its purchasing power (Gabszewicz and Vial, 1972).

3.1.1. Lerner rule in GE. Based on firms' profits (12), we characterize the optimal markups and generalize the Lerner rule to GE as follows.

Lemma 1 (Lerner rule in GE)**.** *Each firm's optimal markup satisfies the Lerner rule*

$$\underbrace{\frac{\mu_i}{\mu_i - 1}}_{\text{inverse of Lerner index}} = \underbrace{1 - \frac{\partial \log R_i}{\partial \log \mu_i}}_{\text{GE demand elasticity}} \tag{13}$$

where $1 - \frac{\partial \log R_i}{\partial \log \mu_i}$ *is the demand elasticity with respect to markup in GE.*

To understand this GE Lerner rule, let us recall the canonical monopoly pricing problem in PE (a special case of Proposition 1). With a given demand function $D(p)$ and a constant marginal cost c, maximizing profit $\pi = (p - c)D(p)$ gives rise to the celebrated Lerner rule

$$\frac{\mu}{\mu - 1} = \underbrace{-\frac{\partial \log D(p)}{\partial \log p}}_{\text{demand elasticity}}.$$

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Comparing (13) to this, we note that $1-\frac{\partial \log R_i}{\partial \log \mu_i}$ is the proper generalization of demand elasticity to GE with competition in markups, defined based on the revenue function $R(\cdot)$. This more general definition nests the PE version based on the demand function $D(\cdot)$.[7]

3.2. *Nested CES network*

While the equilibrium concept does not involve any restriction on the production structure, we focus on a nested CES structure and isoelastic consumption–labour trade-off, which will allow for a separation of each firm's demand elasticity into within-industry substitution and industry demand elasticity, as is common in the literature (Atkeson and Burstein, 2008; Edmond *et al.*, 2015). Further, it gives rise to an empirical test of the theory by exploiting across-industry differences. Nonetheless, I show in Supplementary Appendix D that the insights carry over to an arbitrary firm network structure with arbitrary constant-returns production functions.

The consumer's utility is

$$U(C,L)=\begin{cases}\frac{C^{1-\frac{1}{\epsilon-1}}}{1-\frac{1}{\epsilon-1}}-L, & \epsilon\neq 2\\ \log C-L, & \epsilon=2\end{cases}$$

with $\epsilon>1$.[8] This gives rise to the final demand under optimal consumption–labour choice, with the consumer price index P,

$$C=P^{1-\epsilon}. \tag{14}$$

The final demand implies that if there is only one firm in this economy with a linear production technology using labour, it will face a GE demand elasticity of ϵ according to Lemma 1.

There are N industries indexed by $n=1,\ldots,N$. Within each industry n, there are I_n firms each producing a distinct good indexed by $i_n\in\mathcal{I}_n$, where the subscript indicates the industry it belongs to. Thus the set of firms in the economy is $\mathcal{I}=\cup_{n=1,\ldots,N}\mathcal{I}_n$. Goods produced by firms in the same industry are combined into a CES industry composite n with elasticity ρ,

$$y_n=\left[\sum_{i_n\in\mathcal{I}_n}\left(y_{i_n}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}.$$

The final consumer and firms in other industries only use such industry composites. The consumer consumes a CES bundle with elasticity θ_c,

$$C=\left[\sum_{n=1}^{N}\phi_n c_n^{\frac{\theta_c-1}{\theta_c}}\right]^{\frac{\theta_c}{\theta_c-1}}.$$

7. To see this, if we set the revenue function to $R(\mu)=pD(p)=\mu cD(\mu c)$, we would have $1-\frac{\partial \log R}{\partial \log \mu}=-\frac{\partial \log D}{\partial \log p}$, which is exactly the PE demand elasticity.

8. Competition in markups can be applied to GE beyond the quasi-linear utility case. In that more general setting, when firms consider their pricing, a wealth effect on consumption demand would arise, driven by changes in firms' profits. I abstract from this channel to focus on the effects of the input–output network, as this wealth effect depends on firm size rather than its network position.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Firms in the same industry use the same CES bundle of labour and industry composites in production with elasticity θ_p, but they differ in productivities z_{i_n}. Each firm i_n produces output y_{i_n} according to

$$y_{i_n} = z_{i_n} q_{i_n} \quad \text{with} \quad q_{i_n} = \left[a_{nl} l_{i_n}^{\frac{\theta_p - 1}{\theta_p}} + \sum_{m=1}^{N} a_{nm} x_{i_n m}^{\frac{\theta_p - 1}{\theta_p}} \right]^{\frac{\theta_p}{\theta_p - 1}}$$

in which $l_{i_n}, x_{i_n m}$ denote firm i_n's usage of labour and industry composite m, respectively. As firms minimize input cost, it gives rise to a unit cost of the input bundle for industry n denoted as

$$\mathbf{C}_n \left(\{p_m\}_{m=1,\ldots,N} \right)$$

with CES industry price index $p_m = (\sum_{i_m \in \mathcal{I}_m} p_{i_m}^{1-\rho})^{\frac{1}{1-\rho}}$. Thus, the marginal cost of firm i_n is given by $z_{i_n}^{-1} \mathbf{C}_n(\{p_m\}_{m=1,\ldots,N})$.

The total input bundle, usage of industry composite m and labour by industry n are

$$q_n = \sum_{i_n \in \mathcal{I}_n} q_{i_n}, \quad x_{nm} = \sum_{i_n \in \mathcal{I}_n} x_{i_n m}, \quad l_n = \sum_{i_n \in \mathcal{I}_n} l_{i_n}.$$

The market clears for each industry composite

$$y_n = c_n + \sum_{m=1}^{N} x_{mn}, \quad \forall n = 1, \ldots, N$$

and for labour

$$L = \sum_{n=1}^{N} l_n.$$

3.3. *Firm demand elasticities and markups*

In this nested CES network, firm i_n in industry n earns a profit of

$$\pi_{i_n}^r = \underbrace{\left(1 - \frac{1}{\mu_{i_n}}\right)}_{\text{profit margin}} \underbrace{R_{i_n}}_{\text{firm revenue}} = \underbrace{\left(1 - \frac{1}{\mu_{i_n}}\right)}_{\text{profit margin}} \cdot \underbrace{s_{i_n}}_{\text{market share}} \cdot \underbrace{R_n}_{\text{industry revenue}} \tag{15}$$

in which each firm's revenue $R_{i_n} = \frac{p_{i_n} y_{i_n}}{P}$ is multiplicatively separable in firm i_n's (revenue-based) market share within its industry $s_{i_n} := \frac{p_{i_n} y_{i_n}}{p_n y_n}$ and industry n's revenue $R_n = \frac{p_n y_n}{P}$.

As suggested by Lemma 1, the optimal markup μ_{i_n} hence depends on the elasticities of market share s_{i_n} and industry revenue R_n with respect to μ_{i_n}.

To determine the industry revenue R_n, we introduce the concepts of industry productivity z_n and markup μ_n. Since only industry bundles are used across industries and by the final consumer, z_n and μ_n serve as sufficient statistics summarizing within-industry composition for the purpose of determining the allocation across industries.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

3.3.1. Industry aggregation. Within each industry, the allocation across firms only depends on their markups μ_{i_n} and productivities z_{i_n} since they use the same production technology. Hence the industry markup μ_n and productivity z_n depend on their firm counterparts $\{\mu_{i_n}, z_{i_n}\}_{i_n \in \mathcal{I}_n}$.

Because the industry composite is a CES aggregate of firm outputs, each firm faces a demand of

$$y_{i_n} = \frac{p_{i_n}^{-\rho}}{p_n^{-\rho}} y_n = \frac{\left(\mu_{i_n}/z_{i_n}\right)^{-\rho}}{\left[\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}\right]^{-\frac{\rho}{1-\rho}}} y_n \tag{16}$$

and has a market share of

$$s_{i_n} = \frac{p_{i_n} y_{i_n}}{p_n y_n} = \frac{p_{i_n}^{1-\rho}}{p_n^{1-\rho}} = \frac{\left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}}. \tag{17}$$

We define the industry markup as the ratio of industry revenue to total input cost, which equals the *revenue-weighted* harmonic average of firm markups, as in Edmond *et al.* (2015, 2023) and Grassi (2017),

$$\mu_n := \frac{p_n y_n}{\mathbf{C}_n q_n} = \left[\sum_{i_n \in \mathcal{I}_n} \frac{1}{\mu_{i_n}} s_{i_n}\right]^{-1}. \tag{18}$$

We define the industry productivity as the ratio of output to input bundle used, which is equal to the *output-quantity-weighted* harmonic mean of firm productivities,[9] as in Edmond *et al.* (2015, 2023) and Grassi (2017),

$$z_n := \frac{y_n}{q_n} = \left[\sum_{i_n \in \mathcal{I}_n} \frac{1}{z_{i_n}} \frac{y_{i_n}}{y_n}\right]^{-1}. \tag{19}$$

Together, the industry price relative to the cost of input bundle is $p_n/\mathbf{C}_n = \mu_n/z_n$.

Finally, we define the cost-based market share of firm i_n in industry n as

$$\tilde{s}_{i_n} := \frac{\mathbf{C}_n q_{i_n}}{\mathbf{C}_n q_n} = \frac{\mu_n}{\mu_{i_n}} s_{i_n} \tag{20}$$

which is larger (smaller) than the revenue-based market share s_{i_n} (17) if firm i_n's markup μ_{i_n} is lower (higher) than industry n's markup μ_n.

9. Note that since y_n is a CES bundle of y_{i_n}, the weights $\frac{y_{i_n}}{y_n}$ do not necessarily sum up to one.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Lemma 2. *For firm i_n in industry n, raising its markup changes its market share s_{i_n}, the industry markup μ_n and the industry productivity z_n by*

$$\frac{\partial \log s_{i_n}}{\partial \log \mu_{i_n}} = (1-\rho)\left(1-s_{i_n}\right) \tag{21}$$

$$\frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} = s_{i_n} + \rho\left(\tilde{s}_{i_n} - s_{i_n}\right) \tag{22}$$

$$\frac{\partial \log z_n}{\partial \log \mu_{i_n}} = \rho\left(\tilde{s}_{i_n} - s_{i_n}\right) \tag{23}$$

where $\tilde{s}_{i_n}$ is the cost-based market share of firm i_n.

Regarding the market share, when goods are substitutable within each industry ($\rho > 1$), if firm i_n raises its markup μ_{i_n}, its market share falls proportionally by $(\rho-1)(1-s_{i_n})$, as firms in other industries and the consumer substitute away from good i_n.

Regarding the industry aggregates, we note $\sum_{i_n \in \mathcal{I}_n} \frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} = 1$, $\sum_{i_n \in \mathcal{I}_n} \frac{\partial \log z_n}{\partial \log \mu_{i_n}} = 0$. That is, if all firms within an industry raise markups proportionally, the industry markup goes up with the same proportion but the industry productivity stays the same. The latter occurs because the within-industry allocation remains unchanged.

Further, $\frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} > s_{i_n}$ and $\frac{\partial \log z_n}{\partial \log \mu_{i_n}} > 0$ for firms whose cost-based market shares $\tilde{s}_{i_n}$ are larger than their revenue-based market shares s_{i_n}, which are those that charge markups below the industry markup ($\mu_{i_n} < \mu_n$) according to (20). It means that as low-markup firms raise their markups or high-markup firms lower theirs, the within-industry allocation gets closer to the efficient allocation, leading to a productivity gain.

3.3.2. Lerner index in nested CES networks. Based on the multiplicative decomposition of firms' profits into their profit margin $\left(1-\frac{1}{\mu_{i_n}}\right)$, their market share s_{i_n} and the industry revenue R_n in (15), we can characterize each firm's optimal markup μ_{i_n} as follows.

Proposition 4 (Firm markup)**.** *For firm i_n in industry n, its markup satisfies the Lerner rule*

$$\underbrace{\frac{\mu_{i_n}}{\mu_{i_n}-1}}_{\text{inverse of Lerner index}} = \underbrace{\overbrace{\rho}^{\text{within-industry substitution}} \cdot\left(1-s_{i_n}\right) + \overbrace{\eta_n}^{\text{industry demand elasticity}} \cdot s_{i_n} + \overbrace{\rho\left(\eta_n - \eta_n^z\right)\left(\tilde{s}_{i_n} - s_{i_n}\right)}^{\text{cycle adjustment}}}_{\text{firm demand elasticity}}, \tag{24}$$

with industry demand elasticities with respect to markup and productivity $\eta_n := 1 - \frac{\partial \log R_n}{\partial \log \mu_n}$, $\eta_n^z := 1 - \frac{\partial \log R_n}{\partial \log z_n^{-1}}$.

This proposition states that each firm's demand elasticity is additively separable in three terms. The first term reflects the within-industry substitution among firms, and it is more important for firms with small market shares s_{i_n}. As a firm raises its markup, its market share shrinks as described by (21), more so for smaller firms. The second term represents the impact of each firm's markup on the industry revenue, which is a product of the industry demand elasticity η_n and the firm's market share s_{i_n}. For firms with larger market shares s_{i_n}, raising their markups increases the industry markup more via the first term in (22) and thus affects the industry revenue

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

more. Analogously to the firm-level elasticity $\frac{\partial \log R_{i_n}}{\partial \log \mu_{i_n}}$ of Section 3.1, we define the industry-level elasticity $\frac{\partial \log R_n}{\partial \log \mu_n}$ ($\frac{\partial \log R_n}{\partial \log z_n^{-1}}$) that underlies η_n (η_n^z) as the response of industry revenue when varying industry n's markup (productivity), holding fixed industry n's productivity (markup) and all other industries' markups and productivities, and letting prices and quantities adjust through the second-stage GE. The two conventions are consistent since a firm's markup μ_{i_n} affects industry revenue R_n only through industry markup and productivity (μ_n, z_n). In the next section, I will explicitly characterize η_n, η_n^z in terms of industry network positions, hence microfounding the Lerner index in production networks.

The last term in (24) is an adjustment that reflects the difference between industry demand elasticities η_n and η_n^z with respect to markup and productivity. As the firm raises its markup μ_{i_n}, it moves industry productivity z_n by $\rho(\tilde{s}_{i_n} - s_{i_n})$ and industry markup μ_n by an additional $\rho(\tilde{s}_{i_n} - s_{i_n})$ beyond s_{i_n}, according to (22, 23). If the industry revenue responds differently to an increase in markup μ_n versus a decrease in productivity z_n, *i.e.* $\eta_n \neq \eta_n^z$, this generates an additional effect on firm demand elasticity. I will show shortly that $\eta_n - \eta_n^z$ is positive if industry n is in a cycle and measures the strength of the cycle, and zero otherwise. For industries in cycles, this adjustment term flips sign depending on whether $\tilde{s}_{i_n} > s_{i_n}$, or equivalently, whether firm i_n charges a markup below the industry markup ($\mu_{i_n} < \mu_n$) according to (20). Going forward, I maintain the assumption that $\rho > \eta_n$ so that larger firms within each industry face smaller demand elasticities and charge higher markups. This relation is not disturbed by the adjustment term in our calibration and is consistent with broader empirical evidence.

The pricing formula (24) nests that of Atkeson and Burstein (2008) under simultaneous price-setting in a horizontal economy with a continuum of industries and a finite number of firms within each industry.[10] In that economy, the industry demand elasticity η_n equals the consumer elasticity of substitution θ_c and $\eta_n = \eta_n^z$ holds. Hence (24) simplifies to

$$\frac{\mu_{i_n}}{\mu_{i_n} - 1} = \rho\left(1 - s_{i_n}\right) + \theta_c s_{i_n}.$$

4. CHARACTERIZING INDUSTRY DEMAND ELASTICITIES

In this section, I characterize the industry demand elasticities η_n, η_n^z in terms of each industry's network position by applying the toolkit of Baqaee and Farhi (2020). Specifically, they characterize comparative statics of all prices and quantities in GE with respect to exogenous wedges. It corresponds to the second stage of our definition of GE with competition in markups, where we vary one producer's markup and hold fixed other producers' markups. Proposition 4 suggests that firms' optimal markups depend on industry demand elasticities η_n, η_n^z, which equal one minus the elasticities of industry revenues with respect to their own markup and productivity changes.

I briefly introduce the notation characterizing the across-industry input–output network as in Baqaee and Farhi (2020). Then I explain the industry demand elasticities η_n, η_n^z, which reflect how industry revenues respond to wedges, since Baqaee and Farhi (2020) mainly study how aggregate output responds. I illustrate the industry demand elasticities with three simple examples in Supplementary Appendix C.

10. The equivalence between competition in markups and simultaneous price-setting in Atkeson and Burstein (2008) extends from PE (Proposition 1) to GE since the aggregate price index—and hence final demand—is unresponsive to any individual firm's price in that economy.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

4.1. *Industry-level input–output network definitions*

There are N industries indexed by $n = 1, \ldots, N$. For notational convenience, we include a fictitious firm 0, which exclusively produces the consumption good with a fixed markup $\mu_0 \equiv 1$. We further include the labour factor, which does not use any input to produce, as a fictitious firm L. Then we define input–output matrices of dimension $(1 + N + 1) \times (1 + N + 1)$, where firm 0 is ordered first and firm L last.

The *cost-based input–output matrix* captures the cost share of industry composite m as input for industry n with element $\tilde{\Omega}_{nm} := \frac{p_m x_{nm}}{\mathbf{C}_n q_n}$.[11] Accordingly, the *cost-based Leontief inverse* is defined as $\tilde{\Psi} := (I - \tilde{\Omega})^{-1} = \sum_{k=0}^{\infty} \tilde{\Omega}^k$, which captures both direct and indirect linkages among industries. In particular, we define the *cost-based Domar weight* as $\tilde{\lambda}_n := \tilde{\Psi}_{0n}$ which measures the importance of industry n in final demand.

The *revenue-based input–output matrix* captures the revenue share of industry composite m as input for industry n with element $\Omega_{nm} := \frac{p_m x_{nm}}{p_n y_n} = \frac{\tilde{\Omega}_{nm}}{\mu_n}$. The *revenue-based Leontief inverse* is defined as $\Psi := (I - \Omega)^{-1} = \sum_{k=0}^{\infty} \Omega^k$. Its 0th row corresponds to each industry's *(revenue-based) Domar weight* $\lambda_n := \frac{p_n y_n}{\sum_{m=1}^{N} p_m c_m} = \Psi_{0n}$ since it is the proportion of payment by the consumer that ultimately goes to industry n.

Finally, the *input–output covariance operator* is defined as $\mathrm{Cov}_{\tilde{\Omega}^{(n)}}(\tilde{\Psi}_{(k)}, \Psi_{(l)}) := \sum_{m=0,1,\ldots,N,L} \tilde{\Omega}_{nm} \tilde{\Psi}_{mk} \Psi_{ml} - (\sum_{m=0,1,\ldots,N,L} \tilde{\Omega}_{nm} \tilde{\Psi}_{mk})(\sum_{m=0,1,\ldots,N,L} \tilde{\Omega}_{nm} \Psi_{ml})$, in which $\tilde{\Omega}^{(n)}$ is the nth row of $\tilde{\Omega}$, and $\tilde{\Psi}_{(k)}, \Psi_{(l)}$ are the kth and lth columns of $\tilde{\Psi}, \Psi$, respectively. This is a covariance, since each row of $\tilde{\Omega}$ that represents an industry has non-negative entries which sum up to one.

4.2. *Industry demand elasticities*

Proposition 5 (Industry demand elasticities)**.** *The demand elasticities of industry n with respect to industry markup μ_n and productivity z_n satisfy*

$$\eta_n := 1 - \frac{\partial \log R_n}{\partial \log \mu_n} = \underbrace{cycle_n}_{\text{feedback via cycle}} + \underbrace{(\theta_p - 1)\, SP_n}_{\text{substitution in production}} + \underbrace{(\theta_c - 1)\, SC_n}_{\text{substitution in consumption}} + \underbrace{(\epsilon - 1)\, CP_n}_{\text{consumer price}} + 1 \tag{25}$$

$$\eta_n^z := 1 - \frac{\partial \log R_n}{\partial \log z_n^{-1}} = \eta_n - cycle_n \tag{26}$$

with

$$\begin{aligned} cycle_n &= \Psi_{nn} - 1, \quad SP_n = \sum_{m=1}^{N} \left[\frac{\lambda_m}{\mu_m \lambda_n} Cov_{\tilde{\Omega}^{(m)}} \left(\tilde{\Psi}_{(n)}, \Psi_{(n)} \right) \right], \\ SC_n &= \frac{1}{\lambda_n} Cov_{\tilde{\Omega}^{(0)}} \left(\tilde{\Psi}_{(n)}, \Psi_{(n)} \right), \quad CP_n = \tilde{\lambda}_n. \end{aligned} \tag{27}$$

11. By Shephard's lemma, this is also the elasticity of marginal cost to price $\frac{\partial \log \mathbf{C}_n}{\partial \log p_m}$.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Further, the sum of the two substitution terms is

$$SP_n + SC_n = \tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n. \tag{28}$$

This proposition suggests that the industry demand elasticity with respect to markup, η_n, additively depends on four network statistics, modulated by the elasticity of substitution in production θ_p, elasticity of substitution in consumption θ_c, and final demand elasticity ϵ. These four network statistics are generically non-negative; hence η_n is increasing in $\theta_p, \theta_c, \epsilon$.[12] The monotonicity in θ_p, θ_c suggests that if goods are more substitutable in use, their producers face higher demand elasticities, which is intuitive. The monotonicity in ϵ means that a higher elasticity in final demand leads to higher demand elasticities across the economy. All four network statistics vary across industries. Industry n has a lower demand elasticity η_n if it is in a weaker cycle, induces more substitution in production and consumption (if $\theta_p, \theta_c < 1$), and has a smaller impact on consumer price (if $\epsilon > 1$). In the next section, I will test these systematic predictions.

To understand these results, I start by decomposing the industry revenue

$$R_n = \frac{p_n y_n}{P} = \frac{p_n y_n}{PC} \cdot \underbrace{C}_{\text{final demand (GDP)}} = \underbrace{\lambda_n}_{\text{Domar weight}} \cdot \underbrace{P^{1-\epsilon}}_{\text{final demand}}$$

in which we plug in the definition of Domar weight and final demand (14).[13] Hence the demand elasticity with respect to markup is

$$\eta_n = 1 - \frac{\partial \log R_n}{\partial \log \mu_n} = - \underbrace{\frac{\partial \log \lambda_n}{\partial \log \mu_n}}_{\text{elasticity of Domar weight}} + \underbrace{(\epsilon - 1)\frac{\partial \log P}{\partial \log \mu_n}}_{\text{final demand response}} + 1 \tag{29}$$

and the demand elasticity with respect to productivity $\eta_n^z = 1 - \frac{\partial \log R_n}{\partial \log z_n^{-1}}$ follows a similar decomposition. The Domar weight part captures the response in allocation given GDP, which happens due to changes in relative prices. The final demand response part reflects the change in GDP as the consumer price index P changes. I now analyse these two parts in order, which give rise to the four network statistics CP, cycle, SP, and SC.

Responses in prices (CP term). The response of any industry price p_m to a change in industry n's markup and productivity is

$$\frac{\partial \log p_m}{\partial \log (\mu_n / z_n)} = \tilde{\Psi}_{mn} \tag{30}$$

which implies the response of the aggregate price index (*i.e.* firm 0's price) is $\frac{\partial \log P}{\partial \log(\mu_n / z_n)} = \tilde{\Psi}_{0n} = \tilde{\lambda}_n$. These follow from Shephard's lemma: an increase in industry markup $\mathrm{d} \log \mu_n$ or a decline in productivity $\mathrm{d} \log z_n$ raises its output price $\mathrm{d} \log p_n$ one for one, which directly impacts those industries k who buy from industry n by $\tilde{\Omega}_{kn}$. Since those industries keep their profit margins, their prices adjust by $\tilde{\Omega}_{kn} \mathrm{d} \log p_n$, which triggers further impacts on industries that buy from

12. Theoretically, in the limit of small markups, the cost-based and revenue-based Leontief inverses coincide ($\tilde{\Psi} \approx \Psi$) and hence $\mathrm{Cov}_{\tilde{\Omega}^{(m)}}(\tilde{\Psi}_{(n)}, \Psi_{(n)})$ becomes a variance and is guaranteed non-negative. Empirically, I verify that such a covariance is positive for almost all industry pairs (m, n) in almost all years.

13. As discussed before, we assume that firms maximize real profits and hence we study the real revenue. If we were to assume that firms maximize nominal profits, the final demand term would become a nominal demand $PC = P^{2-\epsilon}$. It would imply that the CP term in (25, 26) will have a coefficient of $\epsilon - 2$ instead of $\epsilon - 1$.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

them. Via many rounds of iterations, the total response of price $\mathrm{d}\log p_m$ to $\mathrm{d}\log(\mu_n/z_n)$ will sum up to $\tilde{\Psi}_{mn}$. In terms of the consumer price P, the more important good n is in final demand (measured by $\tilde{\lambda}_n$), the larger the consumer price response $\frac{\partial \log P}{\partial \log \mu_n}$. This corresponds to the CP term in η_n. Since μ_n/z_n is the sufficient statistic in determining the aggregate price response, the same CP term appears in both η_n and η_n^z.

Responses in Domar weights (cycle, SP, and SC terms). The response of Domar weight λ_n to $\{\mathrm{d}\log \mu_n, \mathrm{d}\log z_n\}$ reflects changes in both prices and quantities. From the definition of input–output network, the Domar weight λ is determined by the revenue-based input–output matrix Ω, whose (m,k) element encodes the share of revenue earned by industry m that is spent on input k. Since each revenue-based Ω_{mk} captures both the cost share $\tilde{\Omega}_{mk}$ and the markup μ_m charged by industry m as $\Omega_{mk} = \frac{\tilde{\Omega}_{mk}}{\mu_m}$, its change is described by

$$\begin{aligned}\mathrm{d}\log \Omega_{mk} &= \underbrace{\mathrm{d}\log \tilde{\Omega}_{mk}\left(\{\mathrm{d}\log p_s\}_{s=1,\ldots,N}\right)}_{\text{substitution due to changes in relative prices}} - \underbrace{\mathrm{d}\log \mu_m}_{\text{change in own markup}} \\ &= \mathrm{d}\log \tilde{\Omega}_{mk}\left(\{\mathrm{d}\log(\mu_s/z_s)\}_{s=1,\ldots,N}\right) - \mathrm{d}\log \mu_m \end{aligned} \tag{31}$$

where the cost share $\tilde{\Omega}$ is a function of all prices, *i.e.* $\tilde{\Omega}_{mk}(\{\mathrm{d}\log p_s\}_{s=1,\ldots,N})$ since firms within each industry minimize cost facing prices of all industry bundles under the same production technology. Further, as shown in the elasticity of industry price (30), changes in prices in response to changes in markups and productivities are summarized by changes in the ratio μ_s/z_s, hence $\tilde{\Omega}_{mk}(\{\mathrm{d}\log(\mu_s/z_s)\}_{s=1,\ldots,N})$. Taken together, $\mathrm{d}\log(\mu_n/z_n)$ affects all elements in Ω through changes in $\tilde{\Omega}$ as all industries and the final consumer respond to changing prices, while $\mathrm{d}\log \mu_n$ additionally changes the nth row of Ω directly through the profit margin.

We first analyse the stand-alone $-\mathrm{d}\log \mu_m$ term in the change in Ω (31), which underlies the cycle term in η_n. Though Ω responds differently to $\mathrm{d}\log \mu_n$ and $\mathrm{d}\log z_n^{-1}$ via this term, this difference affects the Domar weight $\lambda_n = \Psi_{0n}$ only when industry n is in a cycle. To see this, use $\Psi = (I-\Omega)^{-1} = \sum_{k=0}^{\infty} \Omega^k$ to express the Domar weight as

$$\lambda_n = \Psi_{0n} = \sum_{s=0}^{\infty} \sum_{k_1,k_2,\ldots,k_s} \Omega_{0k_1}\Omega_{k_1k_2}\ldots\Omega_{k_sn} \tag{32}$$

which captures the consumer expenditure share on industry composite n—either through the direct purchase Ω_{0n} or through indirect linkages of arbitrary length, such as $\Omega_{0k_1}\Omega_{k_1k_2}\ldots\Omega_{k_sn}$ whereby the consumer buys from industry k_1, which buys from k_2 and so on. The direct effect of $\mathrm{d}\log \mu_n$ on Ω, holding fixed $\tilde{\Omega}$, only changes its nth row, as seen from the change in Ω (31). That is, it only affects the share of revenue industry n spends on other industries ($\Omega_{n\cdot}$), not the revenue shares of other industries buying from industry n ($\Omega_{\cdot n}$). For it to affect the sum in the Domar weight (32), industry n must appear as a buyer in an indirect linkage that ultimately buys from industry n itself. Thus, $\frac{\partial \log \lambda_n}{\partial \log \mu_n} \neq \frac{\partial \log \lambda_n}{\partial \log z_n^{-1}}$ if and only if $\Psi_{nn} > 1$. This is precisely why the cycle term appears in η_n but not in η_n^z. It reflects that when industry n charges a higher markup, less of its revenue goes to upstream industries, which will lower industry n's revenue if those upstreams ultimately buy from n.[14]

14. A limitation of this model is that when firms directly use their own outputs in production, I assume that they buy from themselves and their markups constrain their own input use. It is less of an issue if firms buy from other firms in the same industry or in cycles that involve other industries.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

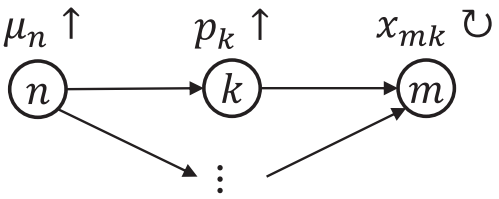


FIGURE 2
Substitution of industry m across all industries $k = 1, \ldots, N$ induced by $\mathrm{d}\log \mu_n$

Finally, we analyse the substitution term in the change in Ω (31) that gives rise to the SP and SC terms in η_n and η_n^z. These capture how relative price changes induced by industry n's markup change cause substitution in other industries' production and the final consumption and thus impact industry n's revenue. The SP term related to production elasticity θ_p deals with the substitution in production by each industry m, illustrated in Figure 2: when industry n raises its markup, it increases each price p_k in the economy by $\tilde{\Psi}_{kn}$. For industry m that sources across all industries k, if inputs are complementary ($\theta_p < 1$), then m substitutes its expenditure towards producers k with higher exposure $\tilde{\Psi}_{kn}$ as their prices become relatively higher. If covariance $\mathrm{Cov}_{\tilde{\Omega}^{(m)}}(\tilde{\Psi}_{(n)}, \Psi_{(n)}) > 0$, it means that the producers that are more exposed to n that m buys from also have high revenue shares that go to n, which is generically the case as discussed in footnote 12. As a result, the substitution of industry m leads to higher revenue of industry n.[15] The SC term related to consumption elasticity θ_c has a similar interpretation, since one can view consumption substitution as the production substitution by the fictitious firm 0 that combines all goods into a consumption bundle.

5. EMPIRICAL ANALYSIS

In this section, I test Propositions 4 and 5 that different industries have different industry demand elasticities η_n, η_n^z, and hence firms with the same market shares in different industries have different markups. I estimate markups following Hall (1986, 1988), De Loecker and Warzynski (2012), and De Ridder *et al.* (2026) using revenue data. De Ridder *et al.* (2026) suggest that this procedure using revenue data recovers the dispersion of markups across firms, but not the average level. Hence I formalize a dispersion-based empirical test that a less elastic industry should see a steeper relation of log markup against log market share. I present evidence in its support based on Compustat firms in the U.S. input–output network from 1997 to 2019.

5.1. *Data*

The empirics feature three sets of variables: firm market shares, firm markups, and industry-level network statistics as determined in Proposition 5. I provide an overview here and leave details of data cleaning and construction in Supplementary Appendix A.

To determine firms' markups and market shares, I use Compustat data from 1997 to 2019. In terms of market shares, for each firm i_n in industry n in year t, its revenue- and cost-based

15. While I assume for simplicity that all industries have the same elasticity of substitution θ_p, one can easily incorporate heterogeneous elasticities θ_m across industries as each industry makes an additive contribution to the SP_n term in η_n, η_n^z.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

market shares are calculated as its sales (cost of goods sold) over the sum of that across all firms in Compustat in the same industry-year

$$\hat{s}_{i_n,t} = \frac{\text{sale}_{i_n,t}}{\sum_{j_n \in \mathcal{I}_{n,t}} \text{sale}_{j_n,t}}, \quad \hat{\tilde{s}}_{i_n,t} = \frac{\text{cogs}_{i_n,t}}{\sum_{j_n \in \mathcal{I}_{n,t}} \text{cogs}_{j_n,t}}. \tag{33}$$

A concern of calculating market shares from Compustat is that it does not include private firms in the same industry. Hence the measured market shares may be overstated relative to the true market shares by an industry-year-specific multiplicative constant. In my empirical strategy using log market shares, this constant will be absorbed by industry-year fixed effects.

To estimate markups, I use the codes from De Ridder *et al.* (2026), which build on Hall (1986, 1988) and De Loecker and Warzynski (2012). I estimate a translog production function with cost of goods sold (the variable input) and capital, for each two-digit NAICS industry, and get estimated markups $\hat{\mu}_{i_n,t}$. As discussed extensively in De Ridder *et al.* (2026), this estimate $\hat{\mu}_{i_n,t}$ using revenue data does not reflect the average level of markups but does inform the dispersion of markups across firms. Following them, I use log markups in my empirical specification.[16]

To determine the industry-level network statistics, I use the U.S. Bureau of Economic Analysis (BEA) input–output table at summary level with seventy-one industries from 1997 to 2019 and keep the sixty-five private-sector industries. I first construct the industry-level input–output matrices and Leontief inverses, $\tilde{\Omega}$, $\tilde{\Psi}$, Ω, Ψ, and then determine the four network statistics (cycle, SP, SC, CP) according to Proposition 5. I further calculate the measure of upstreamness for each industry from Antràs *et al.* (2012), which I control for in the empirical analysis. I assign firms from Compustat to BEA summary industries using their NAICS codes and a concordance provided by the BEA and exclude industries in the commodity, utility, and finance sectors.

5.1.1. Summary statistics. Table 1 presents summary statistics of firm-level and industry-level variables used in my analysis. I include the calibrated industry demand elasticity η, which I will explain in the next subsection. I note that across industries and years, η is uniformly less than one, meaning that the industry revenue would increase with the industry markup. This reflects our calibration of the elasticities of substitution for both production (θ_p) and consumption (θ_c) sufficiently below one.

Figure 3 presents the four network statistics from Proposition 5 as well as the calibrated industry demand elasticity η in 2019 as a heat map. Supplementary Appendix Table B1 lists the top industries by each network statistic in 2019. The order in Figure 3 from top to bottom follows the upstreamness measure from Antràs *et al.* (2012). Industries with the largest cycle statistic are of medium upstreamness, which is intuitive since these industries have to source from other industries while also selling to other industries in a way that closes a loop. The impact on consumer price (CP) statistic is relatively small for most industries. Across industries, upstream ones tend to have larger substitution in production (SP) statistics and smaller substitution in consumption (SC) statistics.[17]

16. De Ridder *et al.* (2026) show that the log of estimated markup correlates strongly with the log of true markups in both simulation and French data.

17. While SP and SC correlate with upstreamness, they are conceptually distinct and can be unrelated in extreme cases. In the exemplar horizontal economy in Supplementary Appendix C, all industries directly sell to consumers and thus they all have upstreamness of one. However, as long as they differ in size, *i.e.* consumption expenditure share $\tilde{\lambda}_n$, they differ in their SC. For another example, in the vertical economy in Supplementary Appendix C, if we take the limit of no labour input for downstream industries, then all industries have the same size $\tilde{\lambda}_n = 1$ and all feature $\text{SP}_n = 0$. However, in that case, they differ in their upstreamness ranging from 1 to N.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

TABLE 1
Summary statistics

	Mean	SD	Min	p25	p50	p75	Max
(a) Firm-level variables							
Sale (millions)	2948	12547	0.0181	86.4	366	1518	450196
Cost of goods sold (millions)	2027	9369	0.0181	47.6	219	969	353261
Estimated market share $\hat{s}$	0.0165	0.0527	2.38×10^{-8}	3.42×10^{-4}	1.74×10^{-3}	9.12×10^{-3}	0.992
Estimated cost share $\hat{\tilde{s}}$	0.0167	0.0537	3.67×10^{-8}	3.26×10^{-4}	1.69×10^{-3}	9.05×10^{-3}	0.98
Estimated markup $\hat{\mu}$	1.52	0.695	0.0767	1.08	1.34	1.86	4.74
Observations	62813						
(b) Industry-level variables							
Cycle	0.101	0.120	0.000	0.010	0.056	0.144	0.674
SP	0.644	0.435	0.000	0.242	0.633	0.990	2.092
SC	0.566	0.353	0.000	0.200	0.586	0.900	1.274
CP	0.043	0.048	0.002	0.013	0.022	0.065	0.315
Calibrated elasticity η	0.596	0.070	0.488	0.550	0.586	0.618	0.967
Upstreamness (ACFH)	1.957	0.729	1.000	1.411	1.840	2.479	4.980
Observations	1495						

Notes: Panel (a) reports summary statistics of Compustat firms from 1997 to 2019. Firm-level sales and cost of goods sold are in millions of 2012 dollars. The revenue- and cost-based market shares $\hat{s}, \hat{\tilde{s}}$ are determined within each BEA summary industry among all Compustat firms after data cleaning, as in (33). The markup is estimated following the production function approach of Hall (1986, 1988), De Loecker and Warzynski (2012), and De Ridder *et al.* (2026). Panel (b) reports summary statistics of network statistics across BEA summary industries from 1997 to 2019. The first four statistics are calculated according to Proposition 5. The calibrated elasticity η is a composite of the first four statistics using empirically estimated weights as in (36). The upstreamness measure follows Antràs *et al.* (2012, ACFH).

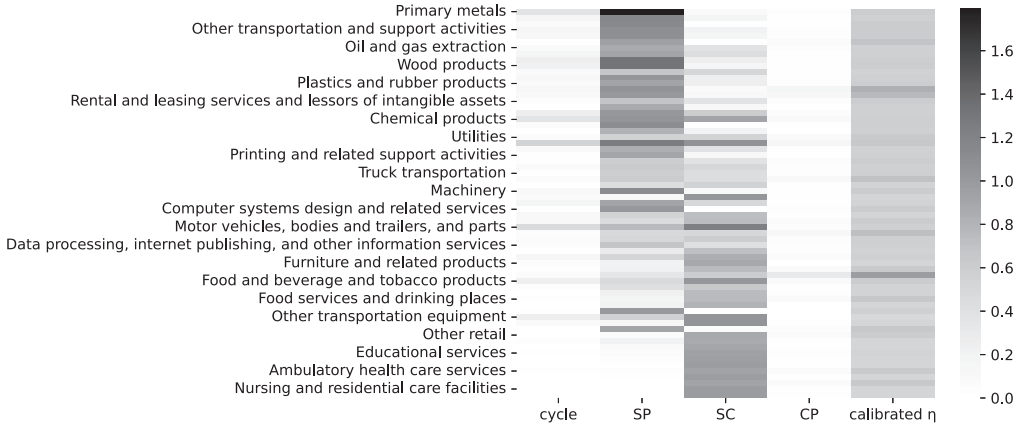


FIGURE 3
Industry-level network statistics in 2019

Notes: The four network statistics are defined in Proposition 5. Industries are sorted from upstream to downstream using the Antràs *et al.* (2012) upstreamness measure. Only selected industry names are presented due to space constraints, but all BEA industries are included in the heat map. Our firm-level empirical analysis excludes commodity, utility, and financial firms.

5.2. *Markup-to-market-share relations across industries*

I test a log–log form of markup-to-market-share relation that is an approximation of Propositions 4 and 5 as follows. While my model is static, I assume it holds year by year.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Proposition 6 (Log–log markup-to-market-share relation)**.** *We have the following log–log relation between the measured markup* $\hat{\mu}_{i_n,t}$ *and market shares* $\hat{s}_{i_n,t}, \hat{\tilde{s}}_{i_n,t}$*, around* $(\bar{\mu}, \bar{s})$*,*

$$\log \hat{\mu}_{i_n,t} \approx \frac{\bar{s}\,(\bar{\mu}-1)^2}{\bar{\mu}}\Bigg[\underbrace{(\rho-\bar{\eta}_t)\log \hat{s}_{i_n,t}}_{\textit{avg effect}} - \underbrace{(\eta_{nt}-\bar{\eta}_t)\log \hat{s}_{i_n,t}}_{\textit{variation across industries}} - \underbrace{\rho\overbrace{(\eta_{nt}-\eta^z_{nt})}^{=cycle_{n,t}}\left(\log \hat{\tilde{s}}_{i_n,t}-\log \hat{s}_{i_n,t}\right)}_{\textit{adjustment}}\Bigg] + const_{nt} \tag{34}$$

in which (η_{nt}, η^z_{nt}) *are the industry demand elasticities characterized in Proposition 5,* $\bar{\eta}_t$ *is the average industry demand elasticity, and* $const_{nt}$ *is an industry-year-specific constant characterized in the proof in Supplementary Appendix F.1.*

This proposition predicts that the markup-to-market-share slope is steeper for an industry with a lower demand elasticity η_{nt}, when controlling for an adjustment term that exists among firms whose cost-based market shares $\hat{\tilde{s}}_{i_n,t}$ differ from the revenue-based market shares $\hat{s}_{i_n,t}$ in industries that are in cycles $(\eta_{nt}-\eta^z_{nt}\neq 0)$.

I run the following panel regression that parallels Proposition 6, plugging in the expressions of (η_{nt}, η^z_{nt}) in terms of network statistics from Proposition 5

$$\begin{aligned}\log \hat{\mu}_{i_n,t} = {} & \underbrace{\left[\gamma_{\text{cycle}}\left(\text{cycle}_{nt}-\overline{\text{cycle}_t}\right)+\gamma_{\text{SP}}\left(\text{SP}_{nt}-\overline{\text{SP}}_t\right)+\gamma_{\text{SC}}\left(\text{SC}_{nt}-\overline{\text{SC}}_t\right)\right.}_{\text{variation across industries}} \\ & \underbrace{\left.+\gamma_{\text{CP}}\left(\text{CP}_{nt}-\overline{\text{CP}}_t\right)\right]\cdot\log \hat{s}_{i_n,t}}_{\text{variation across industries}} \\ & + \underbrace{\beta_t \log \hat{s}_{i_n,t}}_{\text{avg effect}} + \underbrace{\gamma_{\text{adjustment}}\cdot\text{cycle}_{nt}\cdot\left(\log \hat{\tilde{s}}_{i_n,t}-\log \hat{s}_{i_n,t}\right)}_{\text{adjustment}} + \delta_{nt}+\delta_{i_n} \\ & + \underbrace{\left[\gamma_{\text{upstreamness}}\left(\text{upstreamness}_{nt}-\overline{\text{upstreamness}_t}\right)\right]\cdot\log \hat{s}_{i_n,t}}_{\text{control}} + e_{i_n,t}\end{aligned} \tag{35}$$

where the bars denote unweighted means across industries within each year. It is established with a similar markup-on-market-share regression that larger firms charge higher markups (Edmond *et al.*, 2023; Burstein *et al.*, 2025). Building on that, I study the interaction coefficients γ's, which test whether the slope of the markup-to-market-share relation varies across industries in a systematic way predicted by the theory.[18] Propositions 5 and 6 suggest that

$$\gamma_{\text{cycle}}, \gamma_{\text{CP}}, \gamma_{\text{adjustment}} < 0$$

since $\epsilon > 1$. Further, if the elasticities of substitution in production function θ_p and in consumer preference θ_c are less than one, we would have

$$\gamma_{\text{SP}}, \gamma_{\text{SC}} > 0$$

and the reverse if θ_p, θ_c are larger than one.

18. Since firms' sales enter positively into both the markup and market share calculations, measurement errors in sales could create an upward bias in the estimated markup-to-market-share slope. However, this bias, by itself, does not threaten my identification, which relies on the variation of this slope across industries. The network statistics are calculated using the BEA input–output tables, which are a separate data source and should not be affected by any measurement errors in the Compustat data.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

In addition to the industry-year fixed effects δ_{nt} which address mismeasurement of market shares, I control for firm fixed effects δ_{i_n} to purge persistent firm heterogeneity following Edmond *et al.* (2023) and show robustness without firm fixed effects. I allow for a time-varying coefficient β_t of the average effect of market share on markup, which captures potential time variation of the average industry demand elasticity $\bar{\eta}_t$ and the within-industry elasticity of substitution ρ, even though the model holds ρ constant. I further control for an interaction between market share and the upstreamness of the industry from Antràs *et al.* (2012), to address the concern that within-industry elasticity of substitution ρ may vary from upstream to downstream. Standard errors are clustered at the firm level.

Using the empirical estimates, I will calibrate the industry demand elasticity η_{nt} as

$$\eta_{nt} = \text{cycle}_{nt} + \frac{\gamma_{\text{SP}}}{\gamma_{\text{cycle}}}\text{SP}_{nt} + \frac{\gamma_{\text{SC}}}{\gamma_{\text{cycle}}}\text{SC}_{nt} + \frac{\gamma_{\text{CP}}}{\gamma_{\text{cycle}}}\text{CP}_{nt} + 1. \tag{36}$$

5.2.1. Concentration and distortions across industries when demand elasticities vary. Before showing the estimation results of regression (35), I remark on what variable industry demand elasticities imply for the use of concentration as a proxy for markups. My test links markups to market shares and to industry demand elasticities constructed from observed network structures. As markups are challenging to measure directly, concentration alone is often used as an indicator of market power—a practice that is valid in models with constant industry demand elasticities (Grassi, 2017; Burstein *et al.*, 2025; Nocke and Schutz, 2025). I show in Supplementary Appendix E that the relationships of firm markups and industry price indices with concentration can be confounded by variation in industry demand elasticities. In particular, holding fixed the distribution of firm productivities, less elastic industries have lower concentration, higher firm markups, and higher industry price indices. This cautions against naively associating high concentration with high markups.

5.2.2. Estimation results of regression (35). I show the results of regression (35) in Table 2. Column (1) confirms the conventional wisdom that within an industry, markups increase in market shares. Column (2) interacts the log market share with the four network statistics, while controlling for the interaction between the cycle statistic and the log difference between cost-based and revenue-based market shares. All network statistics are highly statistically significant.[19] We note $\gamma_{\text{cycle}}, \gamma_{\text{CP}}, \gamma_{\text{adjustment}} < 0$ as predicted. Further, we find that $\gamma_{\text{SP}}, \gamma_{\text{SC}} > 0$, which suggests that $\theta_p, \theta_c < 1$, namely that different industry composites are complements for both production and consumption.

We use these estimates in column (2) to calculate a calibrated industry demand elasticity η based on (36) and use that as a regressor in column (3). Its coefficient is statistically significant and indistinguishable from that of the cycle interaction term in column (2). Column (4) shows that it is robust to including an interaction between the upstreamness measure and the log market share, which accommodates potential across-industry variation of within-industry elasticity of substitution ρ due to upstreamness. Our calibrated industry demand elasticity η remains significant and the within R^2 barely changes compared to column (3). Column (5) allows for time-varying average effect β_t by including year fixed effects interacted with the log market share, which absorbs time variation in the average industry demand elasticity $\bar{\eta}_t$ as well as any

19. Including the cycle adjustment term mechanically increases the within R^2: since $\log \hat{\tilde{s}}_{i_n,t} - \log \hat{s}_{i_n,t}$ equals the log COGS-to-sale ratio of firm i_n minus that of its industry, and the estimated markup is an output elasticity times the firm's sale-to-COGS ratio, the two variables share the log COGS-to-sale ratio by construction.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

TABLE 2
Markup-to-market-share relations

$\log \hat{\mu}$	(1)	(2)	(3)	(4)	(5)	(6)
$\log \hat{s}$	0.046***	0.040***	0.040***	0.040***		0.021***
	(0.005)	(0.004)	(0.003)	(0.003)		(0.001)
$\log \hat{s} \times$ demeaned calibrated η			−0.232***	−0.245***	−0.246***	−0.222***
			(0.040)	(0.042)	(0.044)	(0.025)
$\log \hat{s} \times$ demeaned cycle		−0.232***				
		(0.069)				
$\log \hat{s} \times$ demeaned SP		0.098***				
		(0.028)				
$\log \hat{s} \times$ demeaned SC		0.116***				
		(0.027)				
$\log \hat{s} \times$ demeaned CP		−0.271***				
		(0.070)				
$\log \hat{s} \times$ demeaned upstreamness				0.005	0.004	
				(0.005)	(0.005)	
$(\log \hat{\tilde{s}} - \log \hat{s}) \times$ cycle		−2.651***	−2.651***	−2.650***	−2.645***	−2.911***
		(0.035)	(0.034)	(0.034)	(0.034)	(0.030)
Firm FE	Yes	Yes	Yes	Yes	Yes	
Industry-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE $\times \log \hat{s}$					Yes	
Observations	62813	62813	62813	62813	62813	62813
Within R^2	0.021	0.519	0.519	0.519	0.520	0.593

Notes: Standard errors in parentheses.
$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.
This table reports estimation results of (35) using Compustat firms from 1997 to 2019. The variables involved are summarized in Table 1. Across all specifications, we control for industry-year fixed effects at the BEA summary industry level, and all columns are estimated on the same sample, fixed at the estimation sample of column (2). Standard errors are clustered at firm level.

potential time variation in the within-industry elasticity of substitution ρ, even though the model holds ρ constant. Finally, column (6) suggests that the findings are robust to not including firm fixed effects.

5.3. *Calibration*

I calibrate the elasticities of substitution θ_p, θ_c across industry composites using estimates from Table 2 column (2), and analyse the calibrated industry demand elasticities. Propositions 5 and 6 suggest that the relative magnitudes of $\gamma_{\text{cycle}}, \gamma_{\text{SP}}, \gamma_{\text{SC}}$ reflect θ_p, θ_c as

$$\theta_p = 1 + \frac{\gamma_{\text{SP}}}{\gamma_{\text{cycle}}}, \quad \theta_c = 1 + \frac{\gamma_{\text{SC}}}{\gamma_{\text{cycle}}}. \tag{37}$$

I then draw a million points from the joint distribution of estimated coefficients $(\gamma_{\text{cycle}}, \gamma_{\text{SP}}, \gamma_{\text{SC}})$ to quantify the uncertainty of the calibration.

Table 3(a) shows that both θ_p and θ_c are significantly below one. While the focus of this paper is not to estimate these parameters, their magnitudes are consistent with estimates in the literature (Atalay, 2017; Chodorow-Reich *et al.*, 2026).

Table 3(b) reports a variance decomposition of the industry demand elasticity η_n across industries in 2019 into the four network statistics, following (36) with the calibrated values of $\theta_p = 0.58, \theta_c = 0.50$, and $\epsilon = 1 + \gamma_{\text{CP}}/\gamma_{\text{cycle}} = 2.17$, the last of which is calibrated by the

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

TABLE 3
Calibration

(a) Calibrated elasticities

Elasticity	Production $\theta_p = 1 + \frac{\gamma_{\text{SP}}}{\gamma_{\text{cycle}}}$	Consumption $\theta_c = 1 + \frac{\gamma_{\text{SC}}}{\gamma_{\text{cycle}}}$
Estimate	0.58	0.50
95% interval	[0.49, 0.64]	[0.28, 0.58]

(b) Variance decomposition of industry demand elasticity η_n in 2019

Component of η_n	SD	Correlation with η_n	Share of Var(η_n)
Cycle_n	0.116	0.043	0.066
$(\theta_p - 1)\,\text{SP}_n$	0.178	−0.220	−0.520
$(\theta_c - 1)\,\text{SC}_n$	0.176	0.318	0.739
$(\epsilon - 1)\,\text{CP}_n$	0.060	0.905	0.714
Sum			1.000

Notes: Panel (a) estimates the elasticity of substitution of CES production and the elasticity of substitution of CES final consumption θ_p, θ_c, defined in Section 3.2. The estimate follows (37). To quantify the uncertainty, I draw a million points from the joint distribution of estimated coefficients from column (2) in Table 2 and calculate θ_p, θ_c according to (37). Their 95% interval reports the 2.5th and 97.5th percentiles of these draws.
Panel (b) reports the variance decomposition of the industry demand elasticity η_n across BEA summary industries in 2019, according to (36). Each row is a coefficient-weighted network statistic, $x_n^j \in \{\text{cycle}_n, (\theta_p - 1)\text{SP}_n, (\theta_c - 1)\text{SC}_n, (\epsilon - 1)\text{CP}_n\}$, so that $\eta_n = 1 + \sum_j x_n^j$ and hence $\text{Var}(\eta_n) = \sum_j \text{Cov}(x_n^j, \eta_n)$. The first two columns report each component's standard deviation and its correlation with η_n. The share of Var(η_n) attributed to each component is its covariance with η_n divided by the variance of η_n; the shares thus sum to one, though individual shares can be negative or exceed one when components are correlated.

same logic as (37). The cycle term makes a weak contribution to the variance of the industry demand elasticity, whereas the SC and CP terms contribute positively and the SP term contributes negatively.

6. AGGREGATE TFP RESPONSES TO FIRM-LEVEL PRODUCTIVITY SHOCKS

Finally, I study the aggregate effects of firm-level productivity shocks when firms endogenously set their markups. To this end, I introduce a productivity shock to one firm and solve for the responses of all firms' markups in my model. Then I characterize the change in the aggregate TFP in response to the productivity shock to one firm and the equilibrium markup responses of all firms. In addition, I compare the TFP response under endogenous markups versus exogenous markups. I show that endogenous markups dampen the TFP response significantly.

6.1. *Solution procedure*

I consider a firm-level productivity shock $\text{d}\log z_{j_n}$ to one firm in industry n. I proceed in two steps: first, I derive a linear system of equations to pin down changes in firm markups $\{\text{d}\log \mu_{i_n}\}$ across all firms, industry markups and productivities $\{\text{d}\log \mu_n, \text{d}\log z_n\}$ across all industries,[20] and industry demand elasticities $\{\text{d}\eta_n, \text{d}\eta_n^z\}$ across all industries. Then I determine the TFP response, which is solely based on the changes in industry markups and productivities $\{\text{d}\log \mu_n, \text{d}\log z_n\}$.

20. Recall that while firm productivities $\{z_{i_n}\}$ are primitive, industry productivities $\{z_n\}$ are equilibrium objects that reflect within-industry allocation.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

6.1.1. Step 1: endogenous markups. In the first step, the linear system can be divided into three blocks about firms, industries, and network, respectively. I summarize the structure of the linear system here and leave the details in Supplementary Appendix F.2. Proposition 4 shows that firm markups μ_{i_n} depend on market shares $s_{i_n}, \tilde{s}_{i_n}$ and the industry demand elasticities η_n, η_n^z. Further, the market shares of each firm $s_{i_n}, \tilde{s}_{i_n}$ as in (17, 20) depend on firm markups and productivities $\{\mu_{j_n}, z_{j_n}\}_{j_n \in \mathcal{I}_n}$ in the same industry n. I linearize and solve these relations to form a linear firm block that pins down firm markup changes in response to changes in productivities and industry demand elasticities

$$\Big(\underbrace{\{\mathrm{d}\log \mu_{i_n}\}_{i_n\in\mathcal{I}_n}}_{\text{firm markups}}\Big) = \mathcal{B}^n_{\text{firm}}\Big(\underbrace{\{\mathrm{d}\log z_{i_n}\}_{i_n\in\mathcal{I}_n}}_{\text{firm productivity shocks}}, \underbrace{\mathrm{d}\eta_n,\ \mathrm{d}\eta_n^z}_{\text{industry demand elasticities}}\Big), \quad \forall n = 1,\dots,N \tag{38}$$

with $\mathcal{B}^n_{\text{firm}}(\cdot)$ being a linear mapping that maps a zero vector to a zero vector for each industry n. This block describes how firms endogenously set markups given the firm-level productivity shocks and industry demand elasticities, taking into account how their market shares adjust.

Because the industry markups $\{\mu_n\}$ (18) and productivities $\{z_n\}$ (19) are weighted averages of firm markups $\{\mu_{i_n}\}$ and productivities $\{z_{j_n}\}$, to the first order, they can be characterized by a linear industry block

$$\Big(\underbrace{\mathrm{d}\log \mu_n,\ \mathrm{d}\log z_n}_{\text{industry markup and productivity}}\Big) = \mathcal{B}^n_{\text{industry}}\Big(\underbrace{\{\mathrm{d}\log \mu_{i_n},\ \mathrm{d}\log z_{i_n}\}_{i_n\in\mathcal{I}_n}}_{\text{firm markups and productivities}}\Big), \quad \forall n = 1,\dots,N. \tag{39}$$

Because the industry demand elasticities η_n, η_n^z depend on industry-level input–output matrices $(\Omega, \tilde{\Omega}, \Psi, \tilde{\Psi})$ as in Proposition 5, and these matrices depend on industry markups and productivities $\{\mu_n, z_n\}_{n=1,\dots,N}$, to the first order, these form a linear network block

$$\Big(\underbrace{\{\mathrm{d}\eta_n,\ \mathrm{d}\eta_n^z\}_{n=1,\dots,N}}_{\text{industry demand elasticities}}\Big) = \mathcal{B}_{\text{network}}\Big(\underbrace{\{\mathrm{d}\log \mu_m,\ \mathrm{d}\log z_m\}_{m=1,\dots,N}}_{\text{industry markups and productivities}}\Big). \tag{40}$$

Together, (38)–(40) determine the responses in firm markups $\{\mu_{i_n}\}$, industry markups and productivities $\{\mu_n, z_n\}$, and industry demand elasticities $\{\eta_n, \eta_n^z\}$ to any firm-level productivity shock.

6.1.2. Step 2: aggregate TFP response. As the across-industry allocation only depends on the industry markups and productivities $\{\mu_n, z_n\}$, the change in aggregate TFP Y/L is, following Baqaee and Farhi (2020) Propositions 2 and 3,

$$\begin{aligned}
\frac{\mathrm{d}\log \mathrm{TFP}}{\mathrm{d}\log z_{i_l}} &= \sum_{k=1,\dots,N}\left(\tilde{\lambda}_k - \tilde{\lambda}_L \sum_{n=0,1,\dots,N} \frac{\lambda_n}{\mu_n}(\theta_n - 1)\,\mathrm{Cov}_{\tilde{\Omega}^{(n)}}\left(\tilde{\Psi}_{(k)}, \frac{\Psi_{(L)}}{\lambda_L}\right)\right)\frac{\mathrm{d}\log z_k}{\mathrm{d}\log z_{i_l}} \\
&\quad + \sum_{k=1,\dots,N}\left(-\tilde{\lambda}_k + \tilde{\lambda}_L\lambda_k\frac{\Psi_{kL}}{\lambda_L} + \tilde{\lambda}_L \sum_{n=0,1,\dots,N} \frac{\lambda_n}{\mu_n}(\theta_n - 1)\,\mathrm{Cov}_{\tilde{\Omega}^{(n)}}\left(\tilde{\Psi}_{(k)}, \frac{\Psi_{(L)}}{\lambda_L}\right)\right)\frac{\mathrm{d}\log \mu_k}{\mathrm{d}\log z_{i_l}}
\end{aligned} \tag{41}$$

where $\theta_0 = \theta_c$, $\mu_0 = 1$ for the final consumer and $\theta_n = \theta_p$, $n = 1,\dots,N$ for all industries.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

6.2. *Application to the U.S. economy*

I apply this procedure to the U.S. input–output network in 2019 and consider the TFP response to a productivity shock to the top firm of each industry. Specifically, I use the input–output matrices constructed from the BEA input–output table as in Section 5 to calculate the four industry-level network statistics that underlie the industry demand elasticities. As before, I exclude commodity, utilities, and finance industries, as well as industries with no or only one firm in Compustat, which leaves forty-eight BEA industries. I set $(\theta_p, \theta_c, \epsilon)$ following the calibration in Section 5.3. Further, I set $\rho = 6$ as the baseline value of the within-industry elasticity of substitution, which implies a markup of $\frac{\rho}{\rho-1} = 1.2$ for the smallest firms, and show robustness with $\rho = 4$ and 10. Then I take the revenue-based and cost-based market shares from Compustat using sales and cost of goods sold to calculate the theory-implied firm-level markups μ_{i_n} based on Proposition 4.[21]

To ensure internal consistency, I use these theory-implied firm markups μ_{i_n}, Compustat measure of revenue-based market shares s_{i_n} and BEA measure of cost-based input–output matrix $\tilde{\Omega}$ to determine the implied cost-based market shares $\tilde{s}_{i_n}$ and the other input–output matrices $\tilde{\Psi}, \Omega, \Psi$.[22] I use these consistent objects in (38)–(41) to solve for the TFP response to a productivity shock to the largest firm of an industry.

We are primarily interested in the TFP response under endogenous markups, as described above. To gauge the importance of endogenous markups, I define the *exogenous markup* benchmark as in Baqaee and Farhi (2020). This benchmark evaluates the TFP response of an economy with the same allocation and firm markups, but assuming that firm markups do not adjust in response to productivity shocks. Specifically, I drop (38), set $\mathrm{d}\log \mu_{i_n} \equiv 0$, and use (39)–(41) to determine the TFP response.

Figure 4 shows the TFP response to a productivity shock to the top firm of each industry. The upper panel shows three bars for each industry. The solid dark and light grey bars represent endogenous and exogenous markups, respectively. The hatched bar shows the size of the firm, measured by its cost-based Domar weight, a relevant measure of the top firm's importance.[23] Across industries, the TFP responses under both endogenous and exogenous markups track the firms' cost-based Domar weights but can be larger or smaller.

The lower panel in Figure 4 shows the TFP response under endogenous markups relative to under exogenous markups. Across industries, endogenous markups dampen transmission by about 6% at the median and up to about 48%. Supplementary Appendix Figure B1 suggests that endogenous markups dampen TFP responses by more for industries that are more concentrated and have lower demand elasticities. This magnitude of dampening is robust to alternative values of the within-industry elasticity of substitution $\rho = 4, 10$ (Figure 4 lower panel) and fairly stable across the past two decades (Supplementary Appendix Figure B2).

To shed light on how endogenous markups dampen the TFP responses to firm-level productivity shocks, Figure 5 plots the responses of firm and industry markups and industry productivities. In the upper panel, the hollow bars with diagonal hatching and with dots show how the top and second largest firms' markups respond to a productivity shock to the top firm.

21. For a small number of firm-year observations whose implied demand elasticity falls below 1.5, so that the implied markup would exceed 3 or be undefined, I cap the markup at 3. No firm is capped in the 2019 baseline; two are at $\rho = 4$, and a few in the other years of Supplementary Appendix Figure B2.

22. One can iterate on firm markups and cost shares to a fixed point where Proposition 4 holds exactly, which leaves the implied aggregate TFP responses essentially unchanged.

23. In a competitive economy with no markups which follows the same cost-based input–output matrix $\tilde{\Omega}$ across industries and cost-based market shares $\tilde{s}$ within industries, this Domar weight $\tilde{s}_{top(l)}\tilde{\lambda}_l$ is the TFP response to firm-level productivity shocks according to Hulten's theorem.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

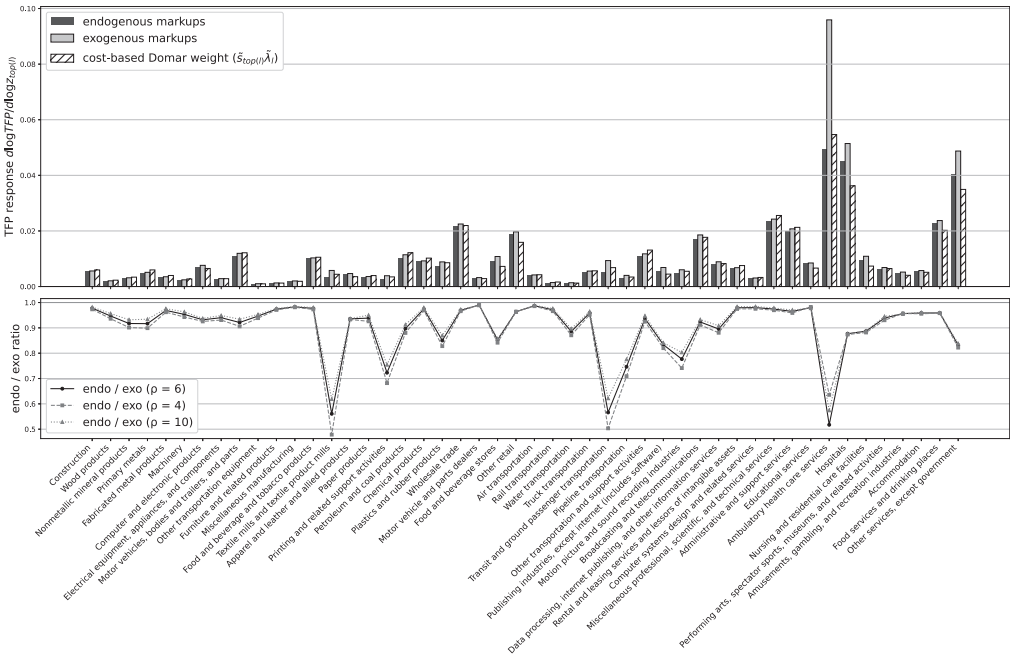


FIGURE 4
TFP response to firm-level productivity shock

Notes: The upper panel shows the TFP response to a productivity shock to the top firm in each industry. Dark grey and light grey bars indicate endogenous markups and exogenous markups, respectively, and the hatched bar shows the cost-based Domar weight of the shocked firm. The lower panel plots the ratio of the TFP response under endogenous markups relative to under exogenous markups: the solid line with circles is the baseline $\rho = 6$, the dashed line with squares $\rho = 4$, and the dotted line with triangles $\rho = 10$.

The dark grey bar indicates the associated industry markup response, and the light grey bar shows the industry markup response under exogenous markups. The lower panel shows the responses of industry productivities under endogenous and exogenous markups. In response to a productivity shock to the top firm, the industry markup rises more and the industry productivity increases less under endogenous markups relative to exogenous markups, leading to a smaller TFP response.

To understand the mechanism, first note that the top firm in each industry is the most productive and has the highest markup. Under exogenous markups, as the top firm gets more productive, it produces more relative to other firms and grows its market share. Thus both industry markup and productivity go up, since they are weighted averages of firm markups and productivities. The industry markup changes purely due to a compositional effect, since firms' markups are held fixed. The economy witnesses an increase in TFP, even though more production is done by the high-markup top firm.

Under endogenous markups, as the top firm gets more productive and grows its market share, it faces a lower demand elasticity and thus charges a higher markup, as seen in the diagonally hatched bar in the upper panel. This increases the industry markup relative to the case with exogenous markups. As the top firm charges a higher markup, it does not grow its production by as much compared to the exogenous-markups case and hence the response of industry productivity is dampened. Meanwhile, the other firms shrink in the industry and lower their markups, suggested by the dotted bar. As firms change their markups, the industry demand elasticities change too, as shown in Supplementary Appendix Figure B3(a), which further influences the within-industry allocation. Firm-level productivity shocks also change industry markups and

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

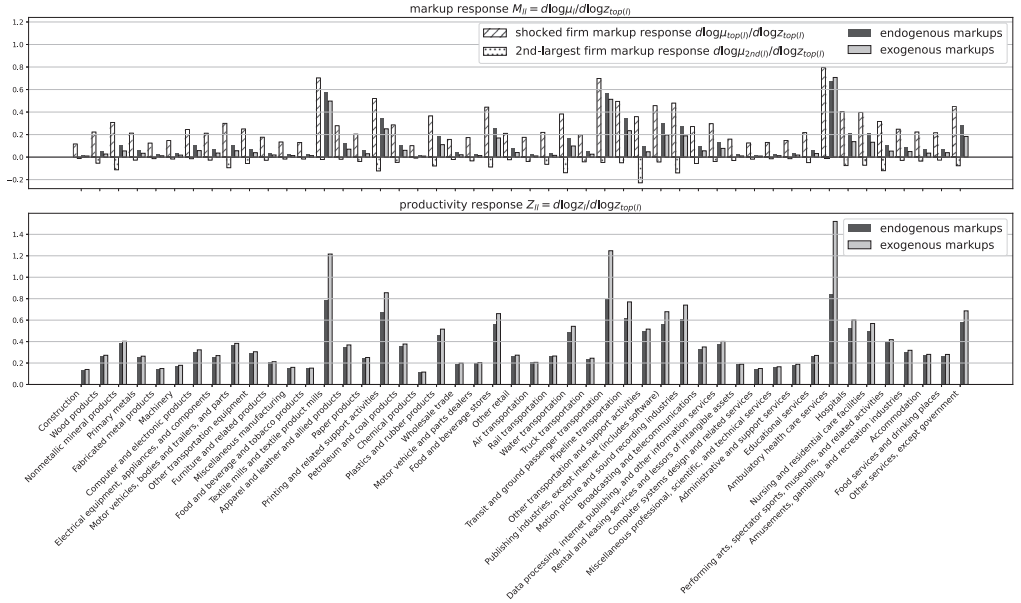


FIGURE 5
Markup and productivity responses to firm-level productivity shocks

Notes: The upper panel shows the responses of firm and industry markups to firm-level productivity shocks to the top firm in each industry. The lower panel plots the responses of industry productivities to firm-level productivity shocks to the top firm in each industry. Dark grey bars: endogenous markups; light grey bars: exogenous markups; hollow bars: firm-level markup responses (diagonal hatching for the shocked top firm, dots for the second-largest firm).

productivities of other industries. Supplementary Appendix Figure B3(b) shows that these cross-industry spillover effects are quantitatively small in the current U.S. economy, which does not imply that they are small in general.

7. CONCLUSION

The celebrated Lerner rule, which links markups inversely to demand elasticities, has inspired decades of research on market power. This research, however, has most often focused on specific industries and typically takes demand functions as given. In contrast, this paper proposes a novel and tractable approach of *competition in markups* to microfound demand elasticities and markups across industries in an arbitrary network economy.

I show that in PE with given final demand functions, competition in markups is equivalent to competition in prices in a horizontal economy with constant marginal costs and is more plausible and tractable in the presence of vertical structures. In GE, firms' markups depend on their market shares and four industry-level network statistics: cycle, substitution in production, substitution in consumption, and consumer-price impact. I confirm these predictions using Compustat data from 1997 to 2019. Endogenous markups significantly dampen the aggregate TFP impact of productivity shocks to top firms in the economy, relative to a benchmark with exogenous markups.

This modelling framework can also help answer many related questions. For example: What are its implications for competition and stabilization policies? How do endogenous markups respond to other wedges in the economy, such as taxes, tariffs, and other trade costs? And how do markups and profitability interact with firm entry and the endogenous formation of production networks?

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Acknowledgments. This paper was previously circulated under the title "Network Lerner Index: Demand and Distortions across Industries". I thank the editor, Thomas Chaney, and three anonymous referees for valuable feedback. I also thank Pol Antràs, David Baqaee, Adrien Bilal, Florin Bilbiie, Justin Bloesch, Gabriel Chodorow-Reich, Emmanuel Farhi, Alessandro Ferrari, Paul Fontanier, Xavier Gabaix, Ben Golub, Basile Grassi, Oliver Hart, Robin Lee, Marc Melitz, Robbie Minton, Lukas Nord (discussant), Thomas Philippon, Elisa Rubbo, Kunal Sangani, Ludwig Straub, Maria Voronina, Alex White, and seminar and conference participants at Harvard and the 2022 UZH PhD Macro Workshop for helpful comments. I thank Zhongji Wei for excellent research assistance and Maarten De Ridder, Basile Grassi, and Giovanni Morzenti for sharing their markup-estimation code.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

Data Availability

The data and code underlying this research are available on Zenodo at https://doi.org/10.5281/zenodo.22837879. The Compustat North America data (S&P Global Market Intelligence, accessed through Wharton Research Data Services) are proprietary and are not included in the replication package; the package provides a script that downloads the extract used in the paper for users with a WRDS subscription. All other data (Bureau of Economic Analysis input–output and price tables, Census Bureau NAICS concordances) are publicly available and are included.

REFERENCES

ACEMOGLU, D., CARVALHO, V. M., OZDAGLAR, A. E., *et al.* (2012), "The Network Origins of Aggregate Fluctuations", *Econometrica*, **80**, 1977–2016.

ACEMOGLU, D. and TAHBAZ-SALEHI, A. (2025), "The Macroeconomics of Supply Chain Disruptions", *Review of Economic Studies*, **92**, 656–695.

ALVIAREZ, V., FIORETTI, M., KIKKAWA, K., *et al.* (2026), "Two-Sided Market Power in Firm-to-Firm Trade" (Working Paper).

AMITI, M., ITSKHOKI, O. and KONINGS, J. (2019), "International Shocks, Variable Markups, and Domestic Prices", *Review of Economic Studies*, **86**, 2356–2402.

ANTRÀS, P., CHOR, D., FALLY, T., *et al.* (2012), "Measuring the Upstreamness of Production and Trade Flows", *American Economic Review: Papers & Proceedings*, **102**, 412–416.

ATALAY, E. (2017), "How Important Are Sectoral Shocks?", *American Economic Journal: Macroeconomics*, **9**, 254–280.

ATKESON, A. and BURSTEIN, A. (2008), "Pricing-to-Market, Trade Costs, and International Relative Prices", *American Economic Review*, **98**, 1998–2031.

AUER, R., BURSTEIN, A. and LEIN, S. M. (2021), "Exchange Rates and Prices: Evidence from the 2015 Swiss Franc Appreciation", *American Economic Review*, **111**, 652–686.

AZAR, J. and VIVES, X. (2021), "General Equilibrium Oligopoly and Ownership Structure", *Econometrica*, **89**, 999–1048.

BAIN, J. S. (1951), "Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940", *The Quarterly Journal of Economics*, **65**, 293–324.

—— (1956), *Barriers to New Competition: Their Character and Consequences in Manufacturing Industries* (Cambridge: Harvard University Press).

BAQAEE, D. R. and FARHI, E. (2019), "The Macroeconomic Impact of Microeconomic Shocks: Beyond Hulten's Theorem", *Econometrica*, **87**, 1155–1203.

—— —— (2020), "Productivity and Misallocation in General Equilibrium", *The Quarterly Journal of Economics*, **135**, 105–163.

BENASSY, J.-P. (1988), "The Objective Demand Curve in General Equilibrium with Price Makers", *The Economic Journal*, **98**, 37–49.

BIGIO, S. and LA'O, J. (2020), "Distortions in Production Networks", *The Quarterly Journal of Economics*, **135**, 2187–2253.

BIZZARRI, M. (2026), "Multilateral Market Power in Input-Output Networks" (Working Paper).

BOWLEY, A. L. (1924), *The Mathematical Groundwork of Economics* (Oxford: Clarendon Press).

BRESNAHAN, T. F. (1989), "Empirical Studies of Industries with Market Power", in *Handbook of Industrial Organization* (Vol. 2) (Amsterdam: North-Holland) 1011–1057.

BULOW, J. I. and PFLEIDERER, P. (1983), "A Note on the Effect of Cost Changes on Prices", *Journal of Political Economy*, **91**, 182–185.

BURSTEIN, A. T., CARVALHO, V. M. and GRASSI, B. (2025), "Bottom-Up Markup Fluctuations", *The Quarterly Journal of Economics*, **140**, 2619–2684.

CAVALLO, A., GOPINATH, G., NEIMAN, B., *et al.* (2021), "Tariff Pass-Through at the Border and at the Store: Evidence from US Trade Policy", *American Economic Review: Insights*, **3**, 19–34.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

CHODOROW-REICH, G., GABAIX, X. and VIVIANO, D. (2026), "Propagation of Shocks in Networks: Identification and Applications" (Working Paper).
DE LOECKER, J. and WARZYNSKI, F. (2012), "Markups and Firm-Level Export Status", *American Economic Review*, **102**, 2437–2471.
DE RIDDER, M., GRASSI, B. and MORZENTI, G. (2026), "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates From Financial Data", *Econometrica*, **94**, 137–168.
DHYNE, E., KIKKAWA, A. K. and MAGERMAN, G. (2022), "Imperfect Competition in Firm-to-Firm Trade", *Journal of the European Economic Association*, **20**, 1933–1970.
EDERER, F. and PELLEGRINO, B. (2026), "A Tale of Two Networks: Common Ownership and Product Market Rivalry", *Review of Economic Studies*, **93**, 1746–1788.
EDMOND, C., MIDRIGAN, V. and XU, D. Y. (2015), "Competition, Markups, and the Gains from International Trade", *American Economic Review*, **105**, 3183–3221.
—— —— —— (2023), "How Costly Are Markups?", *Journal of Political Economy*, **131**, 1619–1675.
FEENSTRA, R. C., GAGNON, J. E. and KNETTER, M. M. (1996), "Market Share and Exchange Rate Pass-Through in World Automobile Trade", *Journal of International Economics*, **40**, 187–207.
FITZROY, F. R. (1974), "Monopolistic Equilibrium, Non-Convexity and Inverse Demand", *Journal of Economic Theory*, **7**, 1–16.
GABAIX, X. (2011), "The Granular Origins of Aggregate Fluctuations", *Econometrica*, **79**, 733–772.
GABSZEWICZ, J. J. and VIAL, J.-P. (1972), "Oligopoly 'A la Cournot' in a General Equilibrium Analysis", *Journal of Economic Theory*, **4**, 381–400.
GANAPATI, S., SHAPIRO, J. S. and WALKER, R. (2020), "Energy Cost Pass-Through in US Manufacturing: Estimates and Implications for Carbon Taxes", *American Economic Journal: Applied Economics*, **12**, 303–342.
GOLDBERG, P. K. and HELLERSTEIN, R. (2013), "A Structural Approach to Identifying the Sources of Local Currency Price Stability", *Review of Economic Studies*, **80**, 175–210.
GRASSI, B. (2017), "IO in I-O: Size, Industrial Organization, and the Input-Output Network Make a Firm Structurally Important" (Working Paper).
HALL, R. E. (1986), "Market Structure and Macroeconomic Fluctuations", *Brookings Papers on Economic Activity*, **1986**, 285–338.
—— (1988), "The Relation between Price and Marginal Cost in U.S. Industry", *Journal of Political Economy*, **96**, 921–947.
HART, O. D. (1985), "Imperfect Competition in General Equilibrium: An Overview of Recent Work", in Arrow, K. and Honkapohja, S. (eds) *Frontiers of Economics* (Oxford: Basil Blackwell) 100–149.
JONES, C. I. (2011), "Intermediate Goods and Weak Links in the Theory of Economic Development", *American Economic Journal: Macroeconomics*, **3**, 1–28.
—— (2013), "Misallocation, Economic Growth, and Input-Output Economics", in Acemoglu, D., Arellano, M. and Dekel, E. (eds) *Advances in Economics and Econometrics* (Cambridge: Cambridge University Press) 419–456.
LEONTIEF, W. W. (1936), "Quantitative Input and Output Relations in the Economic Systems of the United States", *The Review of Economics and Statistics*, **18**, 105–125.
LERNER, A. P. (1934), "The Concept of Monopoly and the Measurement of Monopoly Power", *Review of Economic Studies*, **1**, 157–175.
LIU, E. (2019), "Industrial Policies in Production Networks", *The Quarterly Journal of Economics*, **134**, 1883–1948.
LONG, J. B. and PLOSSER, C. I. (1983), "Real Business Cycles", *Journal of Political Economy*, **91**, 39–69.
MARSCHAK, T. and SELTEN, R. (1974), *General Equilibrium with Price-Making Firms* (Berlin: Springer-Verlag).
MRÁZOVÁ, M. and NEARY, J. P. (2017), "Not So Demanding: Demand Structure and Firm Behavior", *American Economic Review*, **107**, 3835–3874.
NAKAMURA, E. and ZEROM, D. (2010), "Accounting for Incomplete Pass-Through", *Review of Economic Studies*, **77**, 1192–1230.
NEGISHI, T. (1961), "Monopolistic Competition and General Equilibrium", *Review of Economic Studies*, **28**, 196–201.
—— (1972), *General Equilibrium Theory and International Trade* (Amsterdam: North-Holland).
NIKAIDO, H. (1975), *Monopolistic Competition and Effective Demand* (Princeton: Princeton University Press).
NOCKE, V. and SCHUTZ, N. (2025), "Concentration Indices, Welfare Distortions, and Misallocation in Oligopoly" (CEPR Discussion Paper).
PELLEGRINO, B. (2025), "Product Differentiation and Oligopoly: A Network Approach", *American Economic Review*, **115**, 1170–1225.
SALINGER, M. A. (1989), "The Meaning of "Upstream" and "Downstream" and the Implications for Modeling Vertical Mergers", *The Journal of Industrial Economics*, **37**, 373–387.
SANGANI, K. (2025), "Markups Across the Income Distribution: Measurement and Implications" (Working Paper).
SCHMALENSEE, R. (1989), "Inter-Industry Studies of Structure and Performance", in *Handbook of Industrial Organization* (Vol. 2) (Amsterdam: North-Holland) 951–1009.
SPENGLER, J. J. (1950), "Vertical Integration and Antitrust Policy", *Journal of Political Economy*, **58**, 347–352.
TIROLE, J. (1988), *The Theory of Industrial Organization* (Cambridge: The MIT Press).
VORONINA, M. (2022), "Market Power in Production Networks" (Doctoral dissertation, Harvard University Graduate School of Arts and Sciences) 74–101.

Downloaded from academic.oup.com/restud/advance-article/doi/10.1093/restud/rdag123/9031881 by New York University Libraries user on 27 November 2026

Online Appendix for "Competition in Markups:
A Theory of the Lerner Index in Networks and General Equilibrium"

Lingxuan Wu

October 20, 2026

A Data

Firm-level variables. From Compustat North America (S&P Global Market Intelligence, accessed through Wharton Research Data Services), I use the Fundamentals Annual file: consolidated statements (*consol* equal to *C*) in the industrial format (*indfmt* equal to *INDL*) and the standard data format (*datafmt* equal to *STD*), domestic population source, in US dollars, with data dates from 1950-06 to 2022-11, for all companies, with the variables *sic*, *naics*, *ppegt*, *ppent*, *at*, *sale*, *oibdp*, *oiadp*, *emp*, *cogs*, *xsga*, *capx*, *bkvlps*, *mkvalt*, *lt* and the identifying information *conm*, *exchg*, *fic*, *costat*.[1]

Keep firms with *fic* equal to *USA* and *curcd* equal to *USD*, and keep consolidated statements only (*consol* equal to *C*) to avoid double counting.[2] Exclude firms with *sic* between 4900 - 4999 (utility firms which are heavily regulated), between 6000 - 6999 (financial firms whose balance sheets are different), and between 0100 - 1499 (agriculture, forestry, fishing, and mining, including oil and gas extraction; these firms produce undifferentiated commodities and take the world price as given). Exclude firm-year observations with nominal assets (at) less than 10 million; with zero or negative book value, market value, sale, or COGS, or negative employment or capital (*bkvlps*, *mkvalt*, *sale*, *cogs*, *emp*, *ppegt*, *ppent*); or with missing year, SIC, NAICS, sale, COGS, assets, or book liabilities (*fyear*, *sic*, *naics*, *sale*, *cogs*, *at*, *lt*). Deflate sale and COGS using the implicit price deflator for GDP in 2012 dollars. Firms are assigned to BEA industries by matching their NAICS codes in the BEA–NAICS concordance (with housing and other real estate merged, mirroring their treatment in the use tables); firms whose NAICS codes match no BEA industry, or that map to the excluded utilities, finance, and commodity industries (22; 521CI, 523, 524, and 525; 111CA, 113FF, 211, 212, and 213), are dropped.

[1] The replication package provides a script that downloads this extract from the WRDS database.

[2] I also drop one firm (*gvkey* 60972) that is a subsidiary of another Compustat firm (*gvkey* 177108). In the rare cases in which a firm reports two observations for the same fiscal year, I keep the one with the latest report date (*datadate*).

I estimate markups with the published Compustat implementation of De Ridder, Grassi and Morzenti (2026) (version 1.1, September 2025), using their two-factor specification: a translog production function in COGS (the variable input) and capital, estimated separately for each 2-digit NAICS industry. Following their procedure, sales are first purged of measurement error with a second-order polynomial in the two inputs and year fixed effects; the production-function coefficients are then estimated by GMM, with productivity following an AR(1) process with a squared lag, and the lagged variable input and current capital as instruments. Capital is built by the perpetual-inventory method from net property, plant, and equipment. The estimation uses the Compustat panel after the incorporation, currency, and consolidation filters above from 1997 to 2019. Inputs are deflated with the sector deflators distributed with their program, which I extend from 2015 through 2019 using matched BEA gross-output price indexes. Following them, the estimated industry-level elasticity coefficients are merged back onto the full Compustat panel, and the markup is the elasticity, evaluated at the firm's log real COGS, times the firm's sale-to-COGS ratio. For each year, I exclude firms with estimated markups in the top and bottom 5% of the cleaned sample described above.

Industry-level network statistics. I use the BEA use tables before redefinitions in producer value at summary level (71 industries) from 1997 to 2019.[3] I drop the governments, scrap, and noncomparable imports industries and merge housing and other real estate into one. This leaves us with 65 industries. In reading the use tables, I set the small number of negative entries to zero and treat unavailable entries as zero; industries whose commodities are never used as inputs are dropped along with the excluded industries above. The consumption vector is personal consumption expenditures, and the consumer's "revenue" is set equal to its total expenditure.

The use table provides the use of commodities by industry. I make the simplifying assumption that each firm produces one commodity, following Baqaee and Farhi (2020). I construct the revenue-based input-output matrix Ω based on the use of commodities and total industry revenues by each industry, and calculate the revenue-based Leontief inverse $\Psi = (I - \Omega)^{-1}$ accordingly. I construct the cost-based input-output matrix $\tilde{\Omega}$, assuming the total cost is the sum of total intermediate costs and compensation of employees reported in the BEA data, and calculate the cost-based Leontief inverse $\tilde{\Psi} = \left(I - \tilde{\Omega}\right)^{-1}$ accordingly.

[3] Historical series are available too, but the industry definitions vary across years and do not reflect subsequent revisions of NIPAs.

Based on $\Omega, \Psi, \tilde{\Omega}, \tilde{\Psi}$, I construct the determinants of industry elasticities for each industry following the theory predictions in Proposition 5.

Antràs et al. (2012) upstreamness measure. I calculate the measure of industry upstreamness U_n from Antràs et al. (2012) as

$$U_n = 1 + \sum_{m=1}^{N} \frac{x_{mn}}{y_n} U_m = 1 + \sum_{m=1}^{N} \frac{p_n x_{mn}}{p_n y_n} U_m$$

in which x_{mn} is the quantity of industry n's output used by industry m, y_n is the total quantity of industry n's output, and p_n is the price of industry n's output. The second equality lets us read $\frac{p_n x_{mn}}{p_n y_n}$ as dollar shares from the BEA input-output table. We collect these shares into a matrix Δ with $\Delta_{nm} = \frac{p_n x_{mn}}{p_n y_n}$, which implies that the vector of upstreamness U satisfies

$$U = (I - \Delta)^{-1} \mathbf{1} \tag{A1}$$

B Additional Tables and Figures

Table B1: Top industries by each network statistic in 2019

Network statistic	Industry	Value
cycle	Insurance carriers and related activities	0.551
	Motor vehicles, bodies and trailers, and parts	0.466
	Primary metals	0.374
	Chemical products	0.369
	Paper products	0.261
	Other transportation equipment	0.258
	Farms	0.255
	Food and beverage and tobacco products	0.247
SP	Primary metals	1.794
	Wood products	1.322
	Paper products	1.317
	Insurance carriers and related activities	1.268
	Forestry, fishing, and related activities	1.161
	Warehousing and storage	1.147
	Machinery	1.126
	Mining, except oil and gas	1.113
SC	Motor vehicles, bodies and trailers, and parts	1.227
	Insurance carriers and related activities	1.062
	Funds, trusts, and other financial vehicles	1.048
	Other transportation equipment	1.039
	Apparel and leather and allied products	1.031
	Food and beverage and tobacco products	1.025
	Social assistance	0.984
	Nursing and residential care facilities	0.982
CP	Housing and other real estate (merged)	0.315
	Miscellaneous professional, scientific, and technical services	0.170
	Insurance carriers and related activities	0.141
	Administrative and support services	0.128
	Wholesale trade	0.126
	Federal Reserve banks, credit intermediation, and related activities	0.113
	Chemical products	0.098
	Food services and drinking places	0.087

Notes: For each of the four network statistics defined in Proposition 5 (cycle, SP, SC, and CP), the table lists the eight BEA summary industries with the largest values in 2019, computed from the 2019 BEA use table as described in Appendix A. All 65 industries are ranked, including the commodity, utilities, and finance industries that are excluded from the firm-level analysis in Section 5.

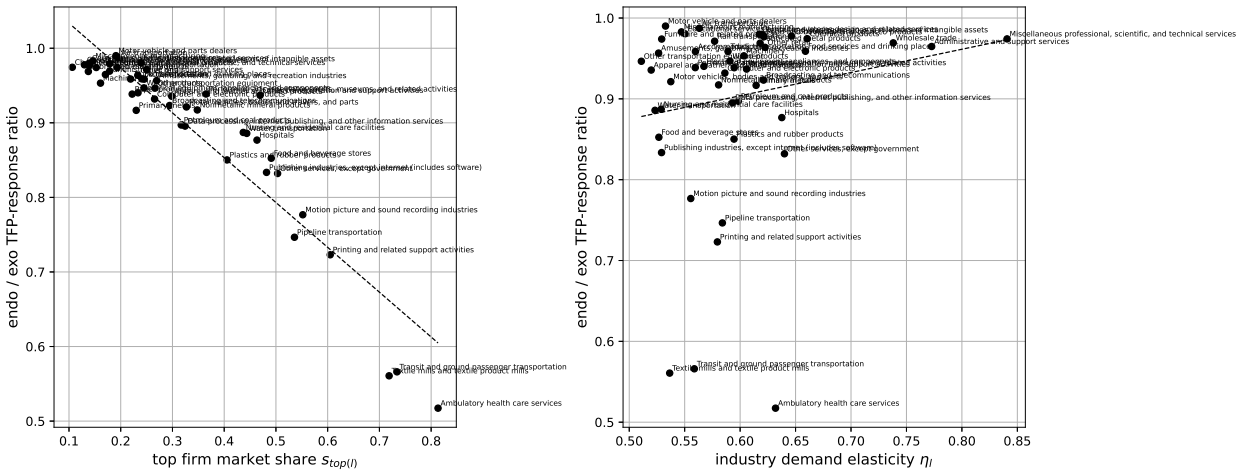


Figure B1: TFP response under endogenous markups relative to exogenous markups against industry concentration and demand elasticity

Notes: Each point is one of the 48 BEA industries in the 2019 baseline with $\rho = 6$. The vertical axis is the ratio of the TFP response to a productivity shock to the industry's top firm, $\mathrm{d}\log TFP/\mathrm{d}\log z_{top(l)}$ in (41), under endogenous markups to the same response under exogenous markups, as in the lower panel of Figure 4; a lower ratio means that endogenous markups dampen the response by more. The left panel plots the ratio against the top firm's revenue-based market share among Compustat firms in its industry, $s_{top(l)}$, defined as in (33), and the right panel against the calibrated industry demand elasticity η_l in (36). Dashed lines are OLS fits.

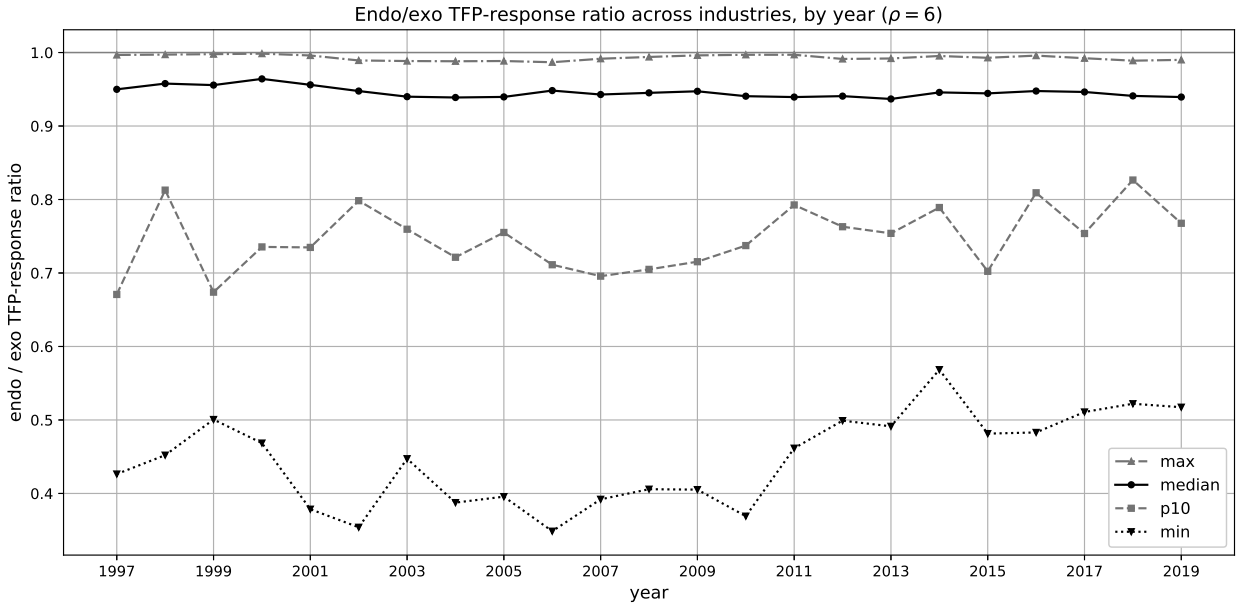
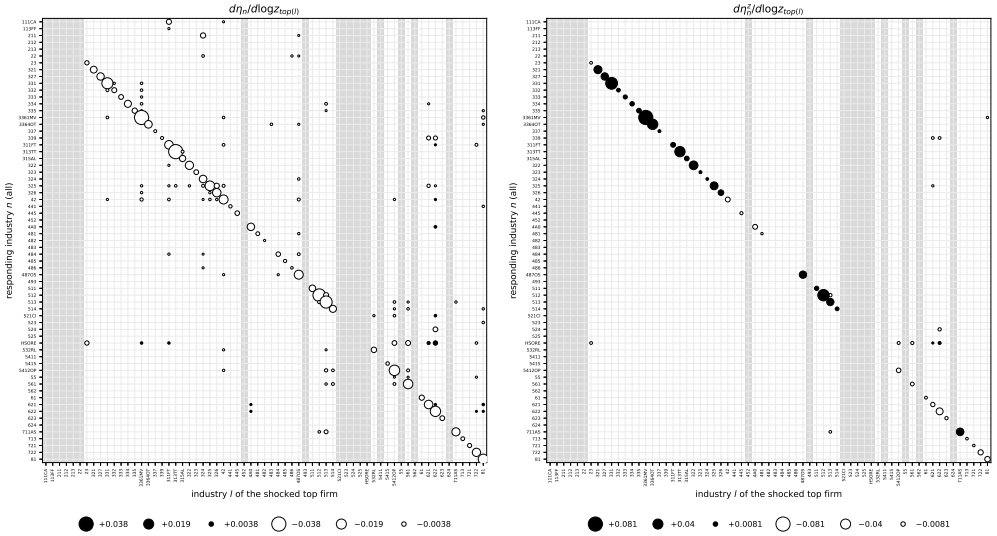
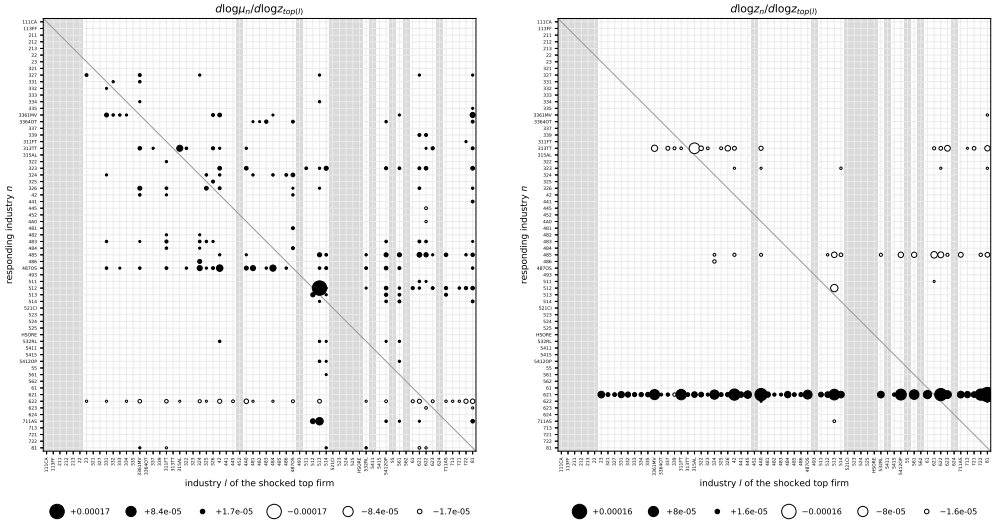


Figure B2: TFP response under endogenous markups relative to exogenous markups

Notes: For each year from 1997 to 2019, the exercise of Section 6 is repeated with that year's BEA input-output table and Compustat market shares, keeping $\left(\theta_p, \theta_c, \epsilon\right)$ at the values calibrated in Section 5.3 and $\rho = 6$. The lines show the maximum, the median, the 10th percentile, and the minimum, across the industries covered in that year (between 48 and 53), of the TFP response to a productivity shock to the industry's top firm under endogenous markups relative to the response under exogenous markups. In sixteen of the years (1997–2008 and 2015–2018), one or two firms have an implied demand elasticity below 1.5 and are assigned a markup of 3, as described before; no firm is capped in 2009–2014 or in 2019.



(a) Industry demand elasticities



(b) Industry markup and productivity

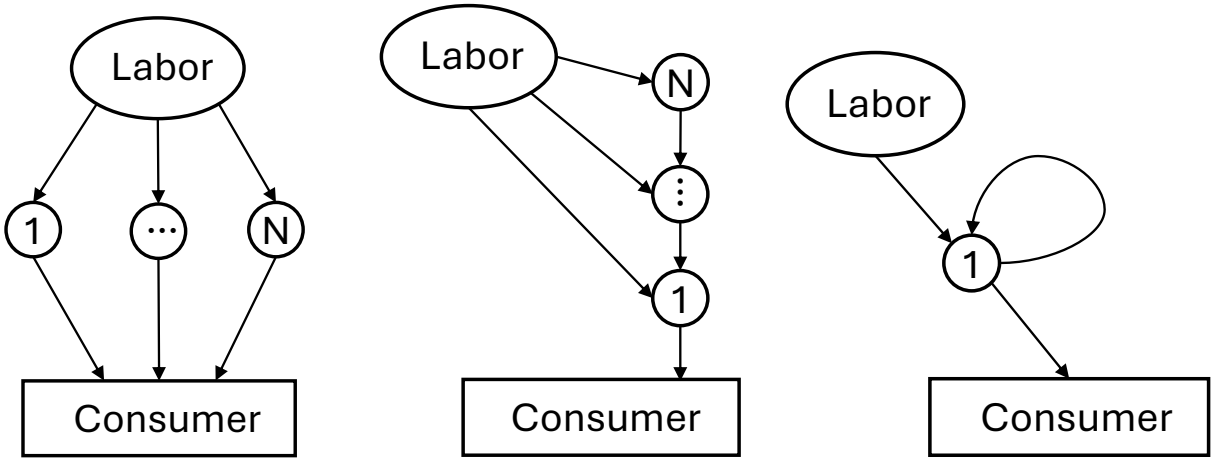
Figure B3: Responses of industry demand elasticities, markup and productivity to firm-level productivity shocks

Notes: Panel (a) shows how industry n's demand elasticities η_n, η_n^z respond to a productivity shock to the top firm in industry l, i.e., $\mathrm{d}\eta_n/\mathrm{d}\log z_{top(l)}$ and $\mathrm{d}\eta_n^z/\mathrm{d}\log z_{top(l)}$. Each circle is one response: filled circles are positive responses, open circles are negative responses, and the area of a circle is proportional to the absolute size of the response, with the scale given below each plot. Responses smaller than 2% of the largest response in a plot are not drawn. Gray columns correspond to industries that are dropped (i.e., the commodity, utilities, and finance sectors, and industries with no or only one firm in Compustat). Panel (b) plots the responses of industry markup and productivity, $\mathrm{d}\log\mu_n/\mathrm{d}\log z_{top(l)}$ and $\mathrm{d}\log z_n/\mathrm{d}\log z_{top(l)}$, for $n \neq l$; the own-industry responses on the diagonal (thin line) are omitted since they are already in Figure 5.

C Examples of Industry Demand Elasticities in General Equilibrium

Here I apply Proposition 5 in a few examples to illustrate each of the four statistics related to network positions in (25, 26).

I consider three examples, illustrated in Figure C1: a horizontal economy, a vertical economy with substitution in production, and a roundabout economy. Compared to the partial equilibrium examples I studied in Section 2, here the demand functions are determined endogenously for all firms in general equilibrium rather than being given as model primitives for downstream firms.



(a) Horizontal (b) Vertical with substitution (c) Roundabout

Figure C1: Examples in general equilibrium

Horizontal economy. In this horizontal economy, illustrated in Figure C1a, firms in each industry n operate a linear production function using labor input. The consumer consumes a CES bundle of all industry composites with elasticity of substitution θ_c.

Applying Proposition 5, the cycle term is zero since the economy is acyclic. Further, an increase in industry n's markup μ_n does not change other industries' input choices, since labor is the only input, and thus the substitution in production term SP is zero too. An increase in industry markup μ_n does trigger a substitution in consumption (SC), as the consumer substitutes towards (away from) other industries if $\theta_c > 1$ ($\theta_c < 1$). Last, an increase in industry markup μ_n affects the final price and thus changes the final

consumption demand. Taking stock, industry n's demand elasticity is

$$\eta_n = \eta_n^z = (\theta_c - 1)\underbrace{\left(1 - \tilde{\lambda}_n\right)}_{\text{SC}} + (\epsilon - 1)\underbrace{\tilde{\lambda}_n}_{\text{CP}} + 1 \tag{C1}$$

where $\tilde{\lambda}_n = \tilde{\Psi}_{0n}$ is the consumption expenditure share on industry n composite in this horizontal economy.

This example also nests Atkeson and Burstein (2008)—if each industry is infinitesimal ($\tilde{\lambda}_n \to 0$), then there is no impact of firms' pricing on the consumer price (CP = 0) and the industry demand elasticity is simply the elasticity of substitution across industry composites in the consumption bundle, i.e., $\eta_n = \theta_c$.

Vertical economy with substitution in production. In this vertical economy with substitution in production, illustrated in Figure C1b, firms in each industry n use a CES production function that combines labor input and the industry-$(n+1)$ composite with elasticity of substitution θ_p—except for firms in the most upstream industry, which use labor input alone. The consumer only consumes the industry-1 composite.

Applying Proposition 5 to this economy, the cycle term is zero too. There is no substitution in consumption associated with an increase in any industry markup μ_n, since the consumer only buys from industry 1. An increase in μ_n leads all downstream industries to substitute towards (away from) labor if $\theta_p > 1$ ($\theta_p < 1$). Last, an increase in industry markup μ_n affects the final price and thus changes the final consumption demand. Taking stock, industry n's demand elasticity is

$$\eta_n = \eta_n^z = \left(\theta_p - 1\right)\underbrace{\left(1 - \tilde{\lambda}_n\right)}_{\text{SP}} + (\epsilon - 1)\underbrace{\tilde{\lambda}_n}_{\text{CP}} + 1 \tag{C2}$$

In this vertical economy, we have $\tilde{\lambda}_N < \cdots < \tilde{\lambda}_1$, since upstream industries have smaller impacts on the consumer price than downstream industries. Thus upstream industries face lower (higher) demand elasticities if $\theta_p < \epsilon$ ($\theta_p > \epsilon$). In the limit with infinitely many industries, industries sufficiently far upstream will have a negligible impact on the consumer price ($\tilde{\lambda}_n \to 0$) and thus face a demand elasticity of θ_p.

Roundabout economy. In a roundabout economy (Figure C1c) where there is only one industry that produces using labor and its own good as intermediate input according to a Cobb-Douglas production technology $y = z\left(\frac{x}{\alpha}\right)^{\alpha}\left(\frac{l}{1-\alpha}\right)^{1-\alpha}$ and charges a markup μ, the industry demand elasticities with respect to markups and productivities are

$$\eta_n = \underbrace{\frac{\alpha}{\mu-\alpha}}_{\text{cycle}} + (\epsilon-1)\underbrace{\frac{1}{1-\alpha}}_{\text{CP}} + 1, \quad \eta_n^z = (\epsilon-1)\underbrace{\frac{1}{1-\alpha}}_{\text{CP}} + 1 \tag{C3}$$

This example illustrates the potential difference between industry demand elasticities with respect to markups and productivities, and the former is strictly higher than the latter in a cyclic economy. In this roundabout economy, a higher markup lowers the share of revenue paid by industry n to intermediate input, which lowers industry n's revenue since it also produces this intermediate input.

D Arbitrary Network Structure and Production Function

While I consider a nested CES network in the main text, I show that the solution concept—general equilibrium with competition in markups—applies to arbitrary network structure and constant-returns production functions. The setup follows Baqaee and Farhi (2020, Appendix H.1).

Firm-level input-output definition and substitution operator. Now I define input-output matrices at the firm level, in a way parallel to the industry-level input-output matrices in Section 4.

There are I firms indexed by $i = 1, \ldots, I$. I include firm 0 which produces the consumption good and the labor factor as a fictitious firm L. Then I define input-output matrices of dimension $(1 + I + 1) \times (1 + I + 1)$, where firm 0 is ordered first and firm L last.

We define the cost-based and revenue-based input-output matrices $\tilde{\Omega}, \Omega$, Leontief inverses $\tilde{\Psi}, \Psi$, and Domar weights $\tilde{\lambda}, \lambda$ at the firm level.

For a firm j with constant-returns cost function $\mathbf{C}_j$, I define the Allen-Uzawa elasticity

of substitution between inputs k and l as

$$\theta_j(k,l) = \frac{\mathbf{C}_j \frac{\partial^2 \mathbf{C}_j}{\partial p_k \partial p_l}}{\frac{\partial \mathbf{C}_j}{\partial p_k}\frac{\partial \mathbf{C}_j}{\partial p_l}} \tag{D1}$$

I define the *input-output substitution operator* for firm j as

$$\Phi_j\left(\tilde{\Psi}_{(l)}, \Psi_{(f)}\right) = -\sum_{i,k} \tilde{\Omega}_{ji}\left(1_{i=k} + \left(\theta_j(i,k) - 1\right)\tilde{\Omega}_{jk}\right)\tilde{\Psi}_{kl}\Psi_{if} \tag{D2}$$

which reduces to $\Phi_j\left(\tilde{\Psi}_{(l)}, \Psi_{(f)}\right) = \left(\theta_j - 1\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(l)}, \Psi_{(f)}\right)$ in the case of CES production function.

Proposition D1. (Firm markup under arbitrary network structure) *For firm i, its markup satisfies the Lerner rule*

$$\frac{\mu_i}{\mu_i - 1} = \overbrace{\underbrace{cycle_i}_{\text{feedback via cycle}} + \underbrace{\mathcal{SP}_i}_{\text{substitution in production}} + \underbrace{\mathcal{SC}_i}_{\text{substitution in consumption}} + \underbrace{(\epsilon - 1)\,CP_i}_{\text{consumer price}}}^{\text{demand elasticity}} + 1 \tag{D3}$$

with

$$\begin{aligned} cycle_i &= \Psi_{ii} - 1, \quad \mathcal{SP}_i = \sum_{j=1}^{I}\left[\frac{\lambda_j}{\mu_j \lambda_i}\Phi_j\left(\tilde{\Psi}_{(i)}, \Psi_{(i)}\right)\right], \\ \mathcal{SC}_i &= \frac{1}{\lambda_i}\Phi_0\left(\tilde{\Psi}_{(i)}, \Psi_{(i)}\right), \quad CP_i = \tilde{\lambda}_i \end{aligned} \tag{D4}$$

This proposition generalizes Propositions 4 and 5 from the nested CES economy to an arbitrary network. The decomposition of demand elasticity into four terms (cycle, substitution in production, substitution in consumption, and final demand response) still holds. The cycle term and the final demand term have the same forms as in the nested CES case. The two substitution terms, which embed the input-output substitution operator, generalize the CES substitution terms to capture non-constant elasticities.

E Concentration and Distortions When Industry Demand Elasticity Varies

A series of papers document rising markups (De Loecker, Eeckhout and Unger, 2020) and rising concentration (Gutiérrez and Philippon, 2017; Autor et al., 2020). Because markups are challenging to measure directly, concentration is sometimes used as a proxy for markups. I illustrate how observed concentration is determined by both firm productivities and variable industry demand elasticities, which may confound the relationship of firm markups and industry price with concentration. In particular, holding fixed firm productivities, a less elastic industry is less concentrated, but has higher firm markups and a higher industry price. That is, a lower concentration means higher markups, against the usual intuition.

To make this point, I study the comparative statics of within-industry allocation and industry price with respect to industry demand elasticity, holding fixed firm productivities. For simplicity, I assume that the industry is not in a cycle so that the adjustment term in Proposition 4 drops out.

Lemma E1. (Sorting within each industry) *For any two firms* $1, 2$ *in the same industry that is not in a cycle, if* 1 *is more productive than* 2*, i.e.,* $z_1 > z_2$*, then at the equilibrium we must have* $\mu_1 > \mu_2, p_1 < p_2, s_1 > s_2$*. That is, within each industry, a more productive firm charges a higher markup and a lower price, and takes a larger market share.*

Lemma E1 suggests that within an industry, the more productive firms take larger market shares and charge higher markups, setting the stage for cross-industry comparison.

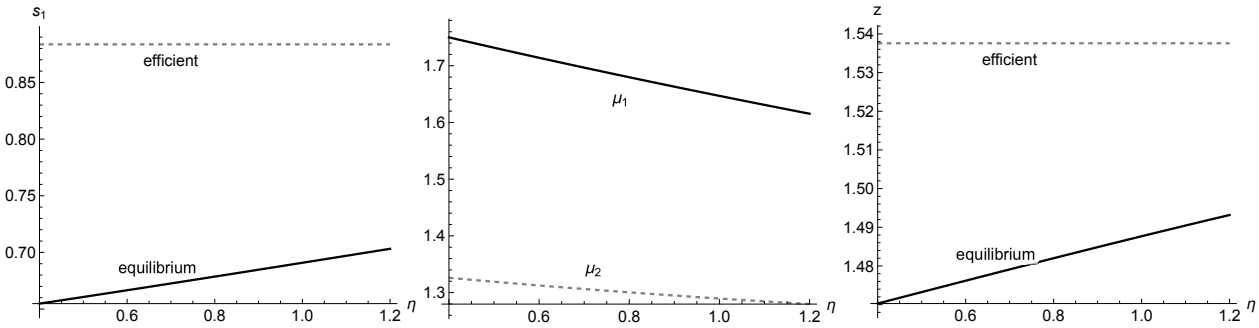
Proposition E1. (Concentration and distortions across industries) Suppose there are only two firms with $z_1 \equiv z > z_2 \equiv 1$ in an industry that is not in a cycle. Define $\kappa := \frac{\log \mu_1 - \log \mu_2}{\log s_1 - \log s_2}$ *as the relation of markup to market share.* We establish the following comparative statics with respect to industry demand elasticity η. As η decreases

1. μ_1, μ_2 *and* $\frac{\mu_1}{\mu_2}$ *increase: in a less elastic industry, firms charge higher markups and the more productive firm charges a particularly high markup relative to the other firm;*

2. s_1 *decreases towards one half and* κ *increases: a less elastic industry is* less *concentrated, and sees a steeper relation of markup to market share;*

3. *industry productivity* z *decreases and price* p *increases: a less elastic industry is* less *productive and has a higher industry price index.*

Proposition E1 presents a series of comparative statics with respect to industry demand elasticity. It suggests that a less elastic industry is less concentrated, but has firms with higher markups and a higher industry price. Figure E1 plots the market allocation, markups, and the industry's productivity, as a function of the industry demand elasticity η.



(a) Larger firm's market share (b) Equilibrium markups (c) Industry productivity

Figure E1: Comparative statics w.r.t. industry demand elasticity η

Notes: Panel (a) shows the equilibrium market share of the more productive firm (firm 1), as well as its hypothetical market share ("efficient") assuming that the two firms charge the same markups, as a function of η. Panel (b) plots two firms' markups. Panel (c) displays the equilibrium industry productivity and the hypothetical industry productivity assuming that the two firms charge the same markups. The figure sets $\rho = 6$, $z_1 = 1.5$, and $z_2 = 1$, and varies η from 0.4 to 1.2

Given Proposition 4, it should not be surprising that the more productive firm charges a higher markup within each industry, and the markup-to-market-share relation is steeper in less elastic industries. The interesting prediction is that given firms' productivity gaps, the less elastic industry has a *lower* concentration, a *lower* productivity and a *higher* price at the industry level. The logic behind concentration is that the more productive firm's demand elasticity depends more on the industry demand elasticity. Thus if the industry demand elasticity is lower, the ratio of the more productive firm's markup to the less productive firm's markup is higher, leading to a smaller market share of the former. As a consequence, if we start with a situation where two firms that have the same productivity and equally split the market in each industry, and increase one firm's productivity, the model predicts that the increase in concentration is smaller in less elastic industries.

Turning to industry-level outcomes, recall that the industry productivity z_n is an output-quantity-weighted harmonic mean of firm productivities, as in (19). Since the productive firm produces relatively less in an industry with a lower demand elasticity,

it contributes less to the industry productivity, resulting in a lower value. The industry price depends on the firms' prices. As both firms set higher markups and higher prices in a less elastic industry, the industry price is higher.

At a high level, it is intuitive that firms in less elastic industries will have higher markups, leading to a higher industry price. In contrast, it can be puzzling at first glance that a less elastic industry has a lower concentration, coexisting with a higher price. Our usual intuition is that in an industry, if the concentration is lower, that means there is more competition and less distortion. That intuition is true if a lower concentration is driven by a larger number of competitors. Indeed, if we fix the industry demand elasticity and populate it with more firms, we will see lower concentration and a lower price. However, concentration is itself an equilibrium outcome—it is a symptom instead of a root cause of imperfect competition. The intuition above overlooks the variable industry demand elasticity as an important driver of concentration. In our model, facing a lower industry demand elasticity, the more productive firm charges an even higher markup relative to the less productive firm, and thus occupies a smaller market share, leading to a lower observed concentration. As the more productive firm contributes less to the industry, the industry becomes less efficient.[4]

F Proofs and Additional Derivations

F.1 Proofs

Proof of Proposition 1. Firm i's profit function is

$$\pi_i = (\mu_i - 1) z_i^{-1} D_i\left(\left\{p_j\right\}_{j\in\{1,\dots,N\}}\right)$$

with $p_i = \mu_i z_i^{-1}$. For firm i's decision, since p_j for $j \neq i$ only depends on μ_j, assuming other firms hold markups fixed amounts to assuming other firms hold prices fixed. Hence, the FOC is

$$0 = \frac{\partial \log \pi_i}{\partial \log \mu_i} = \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log D_i}{\partial \log p_i}$$

[4]This mechanism is reminiscent of the negative relation between concentration and markup from Melitz and Ottaviano (2008), but differs in the root cause. Their mechanism is driven by substitutability among a continuum of heterogeneous firms under free entry. The prediction of this paper stems from strategic interaction among a few firms, and it is driven by the industry elasticity η instead of the within-industry elasticity of substitution ρ.

which yields

$$\frac{\mu_i}{\mu_i - 1} = -\frac{\partial \log D_i}{\partial \log p_i}$$

Two solution concepts give the same equilibrium system, hence the same mapping $\{z_i\} \to (\{\mu_i\}, \{p_i\})$ and the same comparative statics.□

Proof of Proposition 2. Under both assumptions, the downstream firm FOC is

$$0 = \frac{\partial \log \pi_0}{\partial \log \mu_0} = \frac{\mu_0}{\mu_0 - 1} + \frac{\mathrm{d} \log D_0}{\mathrm{d} \log p_0}$$

The downstream firm 0's cost minimization gives rise to a constant marginal cost function $z_0^{-1}\mathbf{C}_0\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right)$ and conditional input demands $d_i\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right)y_0$.

Hence an upstream firm i's profit function is

$$\pi_i = (\mu_i - 1)\, z_i^{-1} d_i\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right) y_0 = (\mu_i - 1)\, z_i^{-1} d_i\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right) D_0\left(p_0\right)$$

with $p_i = \mu_i z_i^{-1}$ for $i \in \{1, \ldots, N\}$ and $p_0 = \mu_0 z_0^{-1}\mathbf{C}_0\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right)$.

SIM-μ. Under SIM-μ, the FOC is

$$\begin{aligned}
0 = \frac{\partial \log \pi_i}{\partial \log \mu_i} &= \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log d_i}{\partial \log p_i}\frac{\partial \log p_i}{\partial \log \mu_i} + \frac{\mathrm{d} \log D_0}{\mathrm{d} \log p_0}\frac{\partial \log p_0}{\partial \log p_i}\frac{\partial \log p_i}{\partial \log \mu_i} \\
&= \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log d_i}{\partial \log p_i} + \frac{\mathrm{d} \log D_0}{\mathrm{d} \log p_0}\frac{\partial \log \mathbf{C}_0}{\partial \log \mu_i} \\
&= \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log d_i}{\partial \log p_i} + \frac{\mathrm{d} \log D_0}{\mathrm{d} \log p_0}\frac{p_i d_i}{\sum_{j\in\{1,\ldots,N\}} p_j d_j}
\end{aligned}$$

where in the last step we use Shephard's lemma. This implies

$$\frac{\mu_i}{\mu_i - 1} = -\frac{\partial \log d_i}{\partial \log p_i} - \frac{\mathrm{d} \log D_0}{\mathrm{d} \log p_0}\frac{p_i d_i}{\sum_{j\in\{1,\ldots,N\}} p_j d_j}$$

where the final demand elasticity $\left(-\frac{\mathrm{d} \log D_0(p_0)}{\mathrm{d} \log p_0}\right)$ plays a role. Further, in response to a downstream productivity shock $\mathrm{d}z_0$, equilibrium p_0 changes and thus the final demand elasticity $\left(-\frac{\mathrm{d} \log D_0(p_0)}{\mathrm{d} \log p_0}\right)$ changes generically, which affects the upstream markups.

SIM-p. Under SIM-p, upstream firm i assumes that downstream price p_0 is fixed and thus the downstream quantity y_0 is fixed. Hence the FOC is

$$0 = \frac{\partial \log \pi_i}{\partial \log \mu_i} = \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log d_i}{\partial \log p_i} \frac{\partial \log p_i}{\partial \log \mu_i} = \frac{\mu_i}{\mu_i - 1} + \frac{\partial \log d_i}{\partial \log p_i}$$

which implies

$$\frac{\mu_i}{\mu_i - 1} = -\frac{\partial \log d_i}{\partial \log p_i}$$

where the final demand elasticity $\left(-\frac{\mathrm{d} \log D_0(p_0)}{\mathrm{d} \log p_0}\right)$ is absent. As a result, in response to a downstream productivity shock $\mathrm{d}z_0$, since the unit demand function $d_i\left(\left\{p_j\right\}_{j\in\{1,\ldots,N\}}\right)$ is unchanged, the upstream markups are unaffected.□

Proof of Proposition 3.

SIM-μ. Under SIM-μ, firm i's demand is

$$y_i = \frac{y_1}{\prod_{j=1}^{i-1} z_j} = \frac{D_1(p_1)}{\prod_{j=1}^{i-1} z_j}$$

with $p_1 = \prod_{i=1}^{N} \mu_i z_i^{-1}$. Firm i's marginal cost is $mc_i = z_i^{-1} \prod_{j=i+1}^{N} \mu_j z_j^{-1}$ and its profit function is $\pi_i = (\mu_i - 1)\, mc_i y_i$. The FOC is

$$0 = \frac{\partial \log \pi_i}{\partial \log \mu_i} = \frac{\mu_i}{\mu_i - 1} + \frac{\mathrm{d} \log D_1(p_1)}{\mathrm{d} \log p_1} \frac{\partial \log p_1}{\partial \log \mu_i} = \frac{\mu_i}{\mu_i - 1} + \frac{\mathrm{d} \log D_1(p_1)}{\mathrm{d} \log p_1}$$

which implies

$$\frac{\mu_i}{\mu_i - 1} = e^{(1)} \tag{F1}$$

with $e^{(1)} = -\frac{\mathrm{d} \log D_1}{\mathrm{d} \log p_1}$. This pricing formula holds for all firms along the vertical chain.

Further, we have, for $j \neq i$

$$\frac{\partial^2 \log \pi_i}{\partial \log \mu_i \partial \log \mu_j} = -\frac{\mathrm{d} e^{(1)}}{\mathrm{d} \log p_1}$$

which is negative—meaning that firms' markup choices are strategic substitutes—if and only if $e^{(1)}$ increases with price p_1.

Now we consider the impact of productivity shocks on downstream price. Since $p_1 = \prod_{i=1}^{N} \mu_i z_i^{-1}$ we have

$$\mathrm{d}\log p_1 = \sum_{i=1}^{N} (\mathrm{d}\log \mu_i - \mathrm{d}\log z_i)$$

Using (F1), we get

$$\mathrm{d}\log \mu_i = -\frac{1}{e^{(1)}-1}\mathrm{d}\log e^{(1)} = -\frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}\mathrm{d}\log p_1,\ \forall i \in \{1,\dots,N\}$$

These form a system of simultaneous equations that jointly determine markups $\{\mathrm{d}\log \mu_i\}_{i\in\{1,\dots,N\}}$ and downstream price $\mathrm{d}\log p_1$. It yields

$$\mathrm{d}\log p_1 = -\frac{\sum_{i=1}^{N} \mathrm{d}\log z_i}{1 + \frac{N}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}}$$

SEQ-p. We solve this iteratively with backward induction, starting with firm 1.

For firm 1, it solves a monopoly pricing problem, taking p_2 as given and chooses a markup

$$\frac{\mu_1}{\mu_1 - 1} = e^{(1)}$$

with $e^{(1)} \equiv -\frac{\mathrm{d}\log D_1}{\mathrm{d}\log p_1}$. Thus firm 1's price satisfies

$$p_1 = \mu_1 z_1^{-1} p_2 = \frac{e^{(1)}}{e^{(1)}-1} z_1^{-1} p_2$$

Taking the total differentiation leads to

$$\mathrm{d}\log p_1 = -\frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}\mathrm{d}\log p_1 - \mathrm{d}\log z_1 + \mathrm{d}\log p_2$$

which implies

$$\mathrm{d}\log p_1 = \frac{-\mathrm{d}\log z_1 + \mathrm{d}\log p_2}{1 + \frac{1}{e^{(1)}-1}\frac{\mathrm{d}\log e^{(1)}}{\mathrm{d}\log p_1}} \tag{F2}$$

as firm 1's local best response.

For firm 2, its demand satisfies $y_2 = \frac{y_1}{z_1} = \frac{D_1(p_1)}{z_1}$. Thus the demand elasticity is

$$e^{(2)} \equiv -\frac{\partial \log y_2}{\partial \log p_2} = -\frac{\mathrm{d} \log D_1}{\mathrm{d} \log p_1}\frac{\partial \log p_1}{\partial \log p_2} = \frac{e^{(1)}}{1 + \frac{1}{e^{(1)}-1}\frac{\mathrm{d} \log e^{(1)}}{\mathrm{d} \log p_1}} \tag{F3}$$

where we plug in (F2). Hence firm 2 charges an optimal markup satisfying

$$\frac{\mu_2}{\mu_2 - 1} = e^{(2)}$$

Note that $e^{(2)}$ only depends on p_1 and reflects the second-order derivative of the final demand function. Firm 2's price satisfies

$$p_2 = \mu_2 z_2^{-1} p_3 = \frac{e^{(2)}}{e^{(2)} - 1} z_2^{-1} p_3$$

Taking the total differentiation yields

$$\mathrm{d} \log p_2 = -\frac{1}{e^{(2)} - 1}\frac{\mathrm{d} \log e^{(2)}}{\mathrm{d} \log p_1}\mathrm{d} \log p_1 - \mathrm{d} \log z_2 + \mathrm{d} \log p_3 \tag{F4}$$

Combining this with (F2) to cancel out $\mathrm{d} \log p_2$, we can determine

$$\mathrm{d} \log p_1^{(2)} = \frac{-\mathrm{d} \log z_1 - \mathrm{d} \log z_2 + \mathrm{d} \log p_3}{1 + \frac{1}{e^{(1)}-1}\frac{\mathrm{d} \log e^{(1)}}{\mathrm{d} \log p_1} + \frac{1}{e^{(2)}-1}\frac{\mathrm{d} \log e^{(2)}}{\mathrm{d} \log p_1}} \tag{F5}$$

This is the response of final price, taking into account both firm 1 and firm 2's best responses. It is also the response of final price in a two-firm vertical economy ($N = 2$) where the upstream firm's input cost is p_3.

We conjecture and verify that firm n charges a markup of

$$\frac{\mu_n}{\mu_n - 1} = e^{(n)}$$

with $e^{(n)}$ defined iteratively as

$$e^{(n)} = \frac{e^{(1)}}{1 + \sum_{i=1}^{n-1} \frac{1}{e^{(i)}-1}\frac{\mathrm{d} \log e^{(i)}}{\mathrm{d} \log p_1}} \tag{F6}$$

and the final price response taking into account firms 1 through n's best responses is

$$\mathrm{d}\log p_1^{(n)} = \frac{-\sum_{i=1}^{n}\mathrm{d}\log z_i + \mathrm{d}\log p_{n+1}}{1+\sum_{i=1}^{n}\frac{1}{e^{(i)}-1}\frac{\mathrm{d}\log e^{(i)}}{\mathrm{d}\log p_1}} \tag{F7}$$

To verify this iteratively, consider the optimal strategy of firm $n+1$. Its demand is $y_{n+1} = \frac{y_1}{\prod_{i=1}^{n} z_i} = \frac{D_1(p_1)}{\prod_{i=1}^{n} z_i}$ and thus the demand elasticity is

$$e^{(n+1)} \equiv -\frac{\partial \log y_{n+1}}{\partial \log p_{n+1}} = -\frac{\mathrm{d}\log D_1}{\mathrm{d}\log p_1}\frac{\partial \log p_1}{\partial \log p_{n+1}} = \frac{e^{(1)}}{1+\sum_{i=1}^{n}\frac{1}{e^{(i)}-1}\frac{\mathrm{d}\log e^{(i)}}{\mathrm{d}\log p_1}} \tag{F8}$$

where we plug in (F7). Thus we verify (F6).

Firm $n+1$'s markup satisfies $\frac{\mu_{n+1}}{\mu_{n+1}-1} = e^{(n+1)}$ and its price hence follows

$$p_{n+1} = \mu_{n+1} z_{n+1}^{-1} p_{n+2} = \frac{e^{(n+1)}}{e^{(n+1)}-1} z_{n+1}^{-1} p_{n+2}$$

Taking total differentiation gives

$$\mathrm{d}\log p_{n+1} = -\frac{1}{e^{(n+1)}-1}\frac{\mathrm{d}\log e^{(n+1)}}{\mathrm{d}\log p_1}\mathrm{d}\log p_1 - \mathrm{d}\log z_{n+1} + \mathrm{d}\log p_{n+2} \tag{F9}$$

Combining this with (F7) to cancel out $\mathrm{d}\log p_{n+1}$, we get

$$\mathrm{d}\log p_1^{(n+1)} = \frac{-\sum_{i=1}^{n+1}\mathrm{d}\log z_i + \mathrm{d}\log p_{n+2}}{1+\sum_{i=1}^{n+1}\frac{1}{e^{(i)}-1}\frac{\mathrm{d}\log e^{(i)}}{\mathrm{d}\log p_1}} \tag{F10}$$

which verifies (F7).

As a result, in the N-firm economy, the final price depends on all productivity shocks as

$$\mathrm{d}\log p_1^{(N)} = \frac{-\sum_{i=1}^{N}\mathrm{d}\log z_i + \mathrm{d}\log p_{N+1}}{1+\sum_{i=1}^{N}\frac{1}{e^{(i)}-1}\frac{\mathrm{d}\log e^{(i)}}{\mathrm{d}\log p_1}}$$

where $\mathrm{d}\log p_{N+1} = 0$ since firm N's input price is the fixed wage.□

Proof of Lemma 1. Taking derivative of firm profits π_i^r (12) w.r.t. markup μ_i yields this result.□

Proof of Lemma 2. The industry composite is $y_n = \left(\sum_{i_n \in \mathcal{I}_n} y_{i_n}^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$ and the industry price is thus $p_n = \left(\sum_{i_n \in \mathcal{I}_n} p_{i_n}^{1-\rho}\right)^{\frac{1}{1-\rho}}$. Each firm faces a demand of

$$y_{i_n} = \frac{p_{i_n}^{-\rho}}{p_n^{-\rho}} y_n = \frac{\left(\mu_{i_n}/z_{i_n}\right)^{-\rho}}{\left[\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}\right]^{-\frac{\rho}{1-\rho}}} y_n$$

and has a market share of

$$s_{i_n} \equiv \frac{p_{i_n} y_{i_n}}{p_n y_n} = \frac{p_{i_n}^{1-\rho}}{p_n^{1-\rho}} = \frac{\left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}}$$

which adds up to one across all firms in the same industry.

We have

$$\frac{\partial \log s_{i_n}}{\partial \log \mu_{j_n}} = (1-\rho) 1_{i_n = j_n} - (1-\rho) \frac{\left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}}{\sum_{k_n \in \mathcal{I}_n} \left(\mu_{k_n}/z_{k_n}\right)^{1-\rho}} = (1-\rho)\left(1_{i_n = j_n} - s_{j_n}\right)$$

We define the within-industry cost-based market share as

$$\tilde{s}_{i_n} \equiv \frac{q_{i_n}}{\sum_{j_n \in \mathcal{I}_n} q_{j_n}} = \frac{q_{i_n} \mathbf{C}_n}{\sum_{j_n \in \mathcal{I}_n} q_{j_n} \mathbf{C}_n} = \frac{\mu_{i_n}^{-1} s_{i_n}}{\sum_{j_n \in \mathcal{I}_n} \mu_{j_n}^{-1} s_{j_n}} = \frac{\mu_{i_n}^{-1} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \mu_{j_n}^{-1} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}}$$

which adds up to one across all firms in the same industry and satisfies

$$\frac{\partial \log \tilde{s}_{i_n}}{\partial \log \mu_{j_n}} = -\rho 1_{i_n = j_n} + \rho \frac{\mu_{j_n}^{-1} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}}{\sum_{k_n \in \mathcal{I}_n} \mu_{k_n}^{-1} \left(\mu_{k_n}/z_{k_n}\right)^{1-\rho}} = -\rho\left(1_{i_n = j_n} - \tilde{s}_{j_n}\right)$$

The industry productivity is defined as the ratio of output to input bundle used

$$z_n = \frac{y_n}{q_n} = \frac{y_n}{\sum_{i_n \in \mathcal{I}_n} q_{i_n}} = \frac{y_n}{\sum_{i_n \in \mathcal{I}_n} z_{i_n}^{-1} y_{i_n}} = \frac{p_n^{-\rho}}{\sum_{i_n \in \mathcal{I}_n} z_{i_n}^{-1} p_{i_n}^{-\rho}} = \frac{\left[\sum_{i_n \in \mathcal{I}_n} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}\right]^{\frac{-\rho}{1-\rho}}}{\sum_{i_n \in \mathcal{I}_n} z_{i_n}^{-1} \left(\mu_{i_n}/z_{i_n}\right)^{-\rho}}$$

The industry markup is defined as the ratio of industry revenue to total input cost

$$\mu_n = \frac{p_n y_n}{\mathbf{C}_n q_n} = \frac{p_n y_n}{\sum_{i_n \in \mathcal{I}_n} \mu_{i_n}^{-1} p_{i_n} y_{i_n}} = \frac{1}{\sum_{i_n \in \mathcal{I}_n} \mu_{i_n}^{-1} s_{i_n}} = \frac{p_n^{1-\rho}}{\sum_{i_n \in \mathcal{I}_n} \mu_{i_n}^{-1} p_{i_n}^{1-\rho}} = \frac{\sum_{i_n \in \mathcal{I}_n} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{i_n \in \mathcal{I}_n} \mu_{i_n}^{-1} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}$$

Hence, we get the impacts of firm's markup on industry aggregates as follows

$$\frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} = (1-\rho) \frac{\left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}} + \rho \frac{\mu_{i_n}^{-1} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \mu_{j_n}^{-1} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}} = (1-\rho) s_{i_n} + \rho \tilde{s}_{i_n}$$

$$\frac{\partial \log z_n}{\partial \log \mu_{i_n}} = -\rho \frac{\left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}} + \rho \frac{\mu_{i_n}^{-1} \left(\mu_{i_n}/z_{i_n}\right)^{1-\rho}}{\sum_{j_n \in \mathcal{I}_n} \mu_{j_n}^{-1} \left(\mu_{j_n}/z_{j_n}\right)^{1-\rho}} = -\rho s_{i_n} + \rho \tilde{s}_{i_n}$$

Further,

$$\frac{\tilde{s}_{i_n}}{s_{i_n}} = \frac{\mu_{i_n}^{-1}}{\sum_{j_n \in \mathcal{I}_n} \mu_{j_n}^{-1} s_{j_n}} = \frac{\mu_n}{\mu_{i_n}}$$

and hence $\frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} > s_{i_n}$ and $\frac{\partial \log z_n}{\partial \log \mu_{i_n}} > 0$ for firms that charge markups below the industry markup, i.e., $\mu_{i_n} < \mu_n$.□

Proof of Proposition 4. The FOC of firm i_n is, based on (15), since R_n depends on μ_{i_n} only through μ_n and z_n,

$$\begin{aligned}\frac{\mu_{i_n}}{\mu_{i_n} - 1} &= -\frac{\partial \log s_{i_n}}{\partial \log \mu_{i_n}} - \frac{\partial \log R_n}{\partial \log \mu_n} \frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} - \frac{\partial \log R_n}{\partial \log z_n} \frac{\partial \log z_n}{\partial \log \mu_{i_n}} + 1 \\ &= -\frac{\partial \log s_{i_n}}{\partial \log \mu_{i_n}} + (\eta_n - 1) \frac{\partial \log \mu_n}{\partial \log \mu_{i_n}} - (\eta_n^z - 1) \frac{\partial \log z_n}{\partial \log \mu_{i_n}} + 1\end{aligned}$$

where

$$\eta_n = 1 - \frac{\partial \log R_n}{\partial \log \mu_n}$$

$$\eta_n^z = 1 - \frac{\partial \log R_n}{\partial \log z_n^{-1}}$$

Using Lemma 2, we can rewrite the FOC as

$$\frac{\mu_{i_n}}{\mu_{i_n} - 1} = -(1-\rho)(1 - s_{i_n}) + (\eta_n - 1)\left[s_{i_n} + \rho\left(\tilde{s}_{i_n} - s_{i_n}\right)\right] - (\eta_n^z - 1)\rho\left(\tilde{s}_{i_n} - s_{i_n}\right) + 1$$

$$= \rho\left(1 - s_{i_n}\right) + \eta_n s_{i_n} + \rho\left(\eta_n - \eta_n^z\right)\left(\tilde{s}_{i_n} - s_{i_n}\right)$$

□

Proof of Proposition 5. The proof largely follows Baqaee and Farhi (2020). They consider an economy with multiple factors in fixed supply whereas I assume as the only factor a variable labor supply under isoelastic consumption-labor trade-off. Due to this difference, it is unnecessary to characterize factor shares in my setting, and I use the wage as numeraire for convenience. Following the decomposition in (29), we first derive the changes in prices and then derive the changes in Domar weight.

Responses in prices. Firms minimize input costs and charge markups. By Shephard's lemma,

$$\mathrm{d}\log p_j = \mathrm{d}\log\left(\mu_j / z_j\right) + \sum_{i=1}^{N} \tilde{\Omega}_{ji} \mathrm{d}\log p_i$$

leading to

$$\mathrm{d}\log p_i = \sum_{j=1}^{N} \tilde{\Psi}_{ij} \mathrm{d}\log\left(\mu_j / z_j\right) \tag{F11}$$

and hence

$$\frac{\partial \log p_i}{\partial \log\left(\mu_j / z_j\right)} = \tilde{\Psi}_{ij} \tag{F12}$$

Setting $i = 0$ gives the change in the aggregate price index $\mathrm{d}\log P$.

Responses in Domar weights. First we derive the changes in input-output matrices $\Omega, \tilde{\Omega}$, and then characterize the changes in Leontief inverses $\Psi, \tilde{\Psi}$. Based on the optimal input choices,

$$\begin{aligned}
\mathrm{d}\Omega_{ji} &= -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1 - \theta_j\right)\tilde{\Omega}_{ji}\left(\mathrm{d}\log p_i - \sum_{k=1}^{N} \tilde{\Omega}_{jk}\mathrm{d}\log p_k\right) \\
&= -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1 - \theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\mathrm{d}\log p, I_{(i)}\right)
\end{aligned}$$

in which $\tilde{\Omega}^{(j)}$ is the j-th row of $\tilde{\Omega}$ and $I_{(i)}$ is the i-th column of identity matrix I and $\theta_0 = \theta_c, \theta_j = \theta_p, \forall j = 1, \ldots, N$. Using (F11), we rewrite $\mathrm{d}\Omega_{ji}$ as

$$\mathrm{d}\Omega_{ji} = -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\left(\mu_k/z_k\right), I_{(i)}\right)$$

Since $\Psi = (I-\Omega)^{-1}$, we have $\mathrm{d}\Psi = \Psi\mathrm{d}\Omega\Psi$ and thus

$$\begin{aligned}
\mathrm{d}\lambda_n = \mathrm{d}\Psi_{0n} = &-\sum_j \Psi_{0j}\mathrm{d}\log\mu_j \sum_i \Omega_{ji}\Psi_{in} \\
&+ \sum_j \frac{\Psi_{0j}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\left(\mu_k/z_k\right), \sum_i I_{(i)}\Psi_{in}\right) \\
&= -\sum_j \Psi_{0j}\mathrm{d}\log\mu_j\left(\Psi_{jn} - 1_{j=n}\right) \\
&+ \sum_j \frac{\Psi_{0j}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\mathrm{d}\log\left(\mu_k/z_k\right), \Psi_{(n)}\right)
\end{aligned}$$

in which the last step uses $\Omega\Psi = \Psi - I$ and $\sum_i I_{(i)}\Psi_{in} = \Psi_{(n)}$. Hence, we have

$$\frac{\partial\log\lambda_n}{\partial\log\mu_n} = -\left(\Psi_{nn}-1\right) + \frac{1}{\lambda_n}\sum_j \frac{\lambda_j}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) \tag{F13}$$

$$\frac{\partial\log\lambda_n}{\partial\log z_n^{-1}} = \frac{1}{\lambda_n}\sum_j \frac{\lambda_j}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) \tag{F14}$$

Plugging these two and (F12) into (29) yields the results in Proposition 5 with

$$\begin{aligned}
\mathrm{cycle}_n &= \Psi_{nn} - 1 \\
\mathrm{SP}_n &= \sum_{m=1}^{N}\left[\frac{\lambda_m}{\mu_m\lambda_n}\mathrm{Cov}_{\tilde{\Omega}^{(m)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right)\right] \\
\mathrm{SC}_n &= \frac{1}{\lambda_n}\mathrm{Cov}_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) \\
\mathrm{CP}_n &= \tilde{\lambda}_n
\end{aligned}$$

Total substitution. Last, to show $\text{SP}_n + \text{SC}_n = \tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n$, we have

$$\begin{aligned}
\text{SP}_n + \text{SC}_n &= \lambda_n^{-1} \sum_j \frac{\lambda_j}{\mu_j} \text{Cov}_{\tilde{\Omega}^{(j)}} \left(\tilde{\Psi}_{(n)}, \Psi_{(n)} \right) \\
&= \lambda_n^{-1} \sum_j \frac{\lambda_j}{\mu_j} \left[\sum_m \tilde{\Omega}_{jm} \tilde{\Psi}_{mn} \Psi_{mn} - \left(\sum_m \tilde{\Omega}_{jm} \tilde{\Psi}_{mn} \right) \left(\sum_m \tilde{\Omega}_{jm} \Psi_{mn} \right) \right] \\
&= \lambda_n^{-1} \sum_{j,m} \Psi_{0j} \Omega_{jm} \tilde{\Psi}_{mn} \Psi_{mn} - \lambda_n^{-1} \sum_j \Psi_{0j} \left(\sum_m \tilde{\Omega}_{jm} \tilde{\Psi}_{mn} \right) \left(\sum_m \Omega_{jm} \Psi_{mn} \right) \\
&= \lambda_n^{-1} \sum_m \left(\Psi_{0m} - 1_{0=m} \right) \tilde{\Psi}_{mn} \Psi_{mn} - \lambda_n^{-1} \sum_j \Psi_{0j} \left(\tilde{\Psi}_{jn} - 1_{j=n} \right) \left(\Psi_{jn} - 1_{j=n} \right) \\
&= -\lambda_n^{-1} \tilde{\Psi}_{0n} \Psi_{0n} + \lambda_n^{-1} \sum_j 1_{j=n} \Psi_{0j} \left(\tilde{\Psi}_{jn} + \Psi_{jn} - 1_{j=n} \right) \\
&= -\tilde{\lambda}_n + \tilde{\Psi}_{nn} + \Psi_{nn} - 1
\end{aligned}$$

as claimed, where the first step uses $\Omega_{jm} = \frac{\tilde{\Omega}_{jm}}{\mu_j}$, the second step uses $\Psi\Omega = \Omega\Psi = \Psi - I, \tilde{\Psi}\tilde{\Omega} = \tilde{\Psi} - I$, and the third step cancels out $\sum_m \Psi_{0m} \tilde{\Psi}_{mn} \Psi_{mn}$ and $\sum_j \Psi_{0j} \tilde{\Psi}_{jn} \Psi_{jn}$.□

Proof of Proposition 6. Time subscript t is suppressed for simplicity. First we log linearize the pricing formula (24) around an expansion point $\mu_{i_n} = \bar{\mu}, s_{i_n} = \tilde{s}_{i_n} = \bar{s}$

$$\begin{aligned}
\frac{\bar{\mu}}{\bar{\mu} - 1} - \frac{\bar{\mu}}{(\bar{\mu} - 1)^2} \left(\log \mu_{i_n} - \log \bar{\mu} \right) =& \rho \left(1 - \bar{s} \right) + \eta_n \bar{s} + \left(\eta_n - \rho \right) \bar{s} \left(\log s_{i_n} - \log \bar{s} \right) \\
&+ \rho \left(\eta_n - \eta_n^z \right) \bar{s} \left(\log \tilde{s}_{i_n} - \log s_{i_n} \right)
\end{aligned}$$

Then I write $\hat{\mu}_{i_n}$ in place of μ_{i_n}[5] and plug in $\hat{s}_{i_n} = \chi_n^s s_{i_n}, \hat{\tilde{s}}_{i_n} = \chi_n^{\tilde{s}} \tilde{s}_{i_n}$ with $\chi_n^s, \chi_n^{\tilde{s}}$ being industry-year-specific constants and get

$$\begin{aligned}
\frac{\bar{\mu}}{\bar{\mu} - 1} - \frac{\bar{\mu}}{(\bar{\mu} - 1)^2} \left(\log \hat{\mu}_{i_n} - \log \bar{\mu} \right) =& \rho \left(1 - \bar{s} \right) + \eta_n \bar{s} + \left(\eta_n - \rho \right) \bar{s} \left(\log \hat{s}_{i_n} - \log \chi_n^s - \log \bar{s} \right) \\
&+ \rho \left(\eta_n - \eta_n^z \right) \bar{s} \left(\log \hat{\tilde{s}}_{i_n} - \log \hat{s}_{i_n} - \log \chi_n^{\tilde{s}} + \log \chi_n^s \right)
\end{aligned}$$

[5] The estimated markups inform the dispersion but not the average level of the true markups. Here I use $\hat{\mu}_{i_n}$ in place of μ_{i_n} for simplicity.

or equivalently,

$$\log \hat{\mu}_{i_n} = \frac{(\bar{\mu}-1)^2}{\bar{\mu}} \left[(\rho - \eta_n)\, \bar{s} \log \hat{s}_{i_n} - \rho\, (\eta_n - \eta_n^z)\, \bar{s} \left(\log \hat{\tilde{s}}_{i_n} - \log \hat{s}_{i_n} \right) \right] + \text{const}_n$$

with

$$\begin{aligned} \text{const}_n = {} & (\bar{\mu}-1) + \log \bar{\mu} \\ & - \frac{(\bar{\mu}-1)^2}{\bar{\mu}} \left[\rho\, (1-\bar{s}) + \eta_n \bar{s} + (\rho - \eta_n)\, \bar{s}\, (\log \chi_n^s + \log \bar{s}) - \rho\, (\eta_n - \eta_n^z)\, \bar{s} \left(\log \chi_n^{\tilde{s}} - \log \chi_n^s \right) \right] \end{aligned}$$

□

Derivation of GE demand elasticities (C1) in horizontal economy. We have $\text{cycle}_n = 0$ since this economy is acyclic. Further $\text{SP}_n = 0$ since labor is the only factor of production. Using (28), we have $\text{SC}_n = 1 - \tilde{\lambda}_n$.□

Derivation of GE demand elasticities (C2) in vertical economy with substitution in production. We have $\text{cycle}_n = 0$ since this economy is acyclic. Further $\text{SC}_n = 0$ since the consumer only consumes industry-1 composite. Using (28), we have $\text{SP}_n = 1 - \tilde{\lambda}_n$.□

Derivation of GE demand elasticities (C3) in roundabout economy. We completely characterize the prices and quantities in this economy and verify our general formula in Proposition 5.

Since $y = z \left(\frac{x}{\alpha}\right)^{\alpha} \left(\frac{l}{1-\alpha}\right)^{1-\alpha}$, the unit cost of input bundle is

$$\mathbf{C} = p^{\alpha}$$

Thus the price $p = \mu z^{-1} \mathbf{C}$ simplifies to

$$p = \left(\frac{\mu}{z}\right)^{\frac{1}{1-\alpha}}$$

The final consumption (14) is

$$C = p^{1-\epsilon} = \left(\frac{\mu}{z}\right)^{\frac{1-\epsilon}{1-\alpha}}$$

Using the market clearing condition $y = x + C$ and the input demand $px = \alpha \frac{py}{\mu}$, we get

$$y = \frac{C}{1 - \frac{\alpha}{\mu}} = \frac{\left(\frac{\mu}{z}\right)^{\frac{1-\epsilon}{1-\alpha}}}{1 - \frac{\alpha}{\mu}}$$

Hence, the demand elasticities are, using $R = \frac{py}{p} = y$

$$\eta = 1 - \frac{\partial \log R}{\partial \log \mu} = 1 - \frac{\partial \log y}{\partial \log \mu} = \frac{\epsilon - \alpha}{1 - \alpha} + \frac{\alpha}{\mu - \alpha}$$
$$\eta^z = 1 - \frac{\partial \log R}{\partial \log z^{-1}} = 1 - \frac{\partial \log y}{\partial \log z^{-1}} = \frac{\epsilon - \alpha}{1 - \alpha}$$

We verify that our general formula in Proposition 5 applies here since $\theta_p = \theta_c = 1$, $\Psi_{nn} = \left(1 - \frac{\alpha}{\mu}\right)^{-1} = \frac{\mu}{\mu - \alpha}$, $\tilde{\lambda} = \frac{1}{1-\alpha}$.□

Proof of Proposition D1. Similar to the nested CES case, we have firm revenue

$$R_i = \frac{p_i y_i}{P} = \frac{p_i y_i}{PC} \cdot C = \lambda_i P^{1-\epsilon}$$

and thus the optimal markup satisfies

$$\frac{\mu_i}{\mu_i - 1} = 1 - \frac{\partial \log R_i}{\partial \log \mu_i} = -\frac{\partial \log \lambda_i}{\partial \log \mu_i} + (\epsilon - 1) \frac{\partial \log P}{\partial \log \mu_i} + 1 \tag{F15}$$

Responses in prices to changes in markups are the same as (F12), since only Shephard's lemma is used regardless of the functional form of the production function

$$\frac{\partial \log p_i}{\partial \log \mu_j} = \tilde{\Psi}_{ij} \tag{F16}$$

Now we characterize the responses in Domar weight. Let x_{jl} denote the quantity of input l used by firm j. By Shephard's lemma, we have

$$x_{jl} = \frac{\partial \mathbf{C}_j}{\partial p_l} y_j$$

Hence the elasticity of demand l w.r.t. price k is

$$\frac{\partial \log x_{jl}}{\partial \log p_k} = \frac{\partial^2 \mathbf{C}_j}{\partial p_l \partial p_k} \frac{p_k}{x_{jl}} y_j = \theta_j(k,l) \frac{\frac{\partial \mathbf{C}_j}{\partial p_k} \frac{\partial \mathbf{C}_j}{\partial p_l}}{\mathbf{C}_j} \frac{p_k}{x_{jl}} y_j = \theta_j(k,l) \frac{p_k x_{jk}}{\mathbf{C}_j y_j} = \theta_j(l,k) \tilde{\Omega}_{jk}$$

where in the second-to-last step we use Shephard's lemma too.

Hence, in response to all price changes, the change in input demand is

$$\mathrm{d} \log x_{jl} = \sum_k \theta_j(l,k) \tilde{\Omega}_{jk} \mathrm{d} \log p_k$$

Since the cost function is homogeneous of degree one, the input demand x_{jl} is homogeneous of degree zero by Shephard's lemma, which means

$$\sum_k \theta_j(l,k) \tilde{\Omega}_{jk} = 0 \tag{F17}$$

Using $\tilde{\Omega}_{ji} = \frac{p_i x_{ji}}{\sum_k p_k x_{jk}}$ we can characterize the change in the cost-based input-output matrix as

$$\begin{aligned}
\mathrm{d} \log \tilde{\Omega}_{ji} &= \mathrm{d} \log p_i + \mathrm{d} \log x_{ji} - \sum_k \tilde{\Omega}_{jk} \left(\mathrm{d} \log p_k + \mathrm{d} \log x_{jk} \right) \\
&= \mathrm{d} \log p_i + \sum_l \theta_j(i,l) \tilde{\Omega}_{jl} \mathrm{d} \log p_l - \sum_k \tilde{\Omega}_{jk} \mathrm{d} \log p_k - \sum_k \tilde{\Omega}_{jk} \sum_l \theta_j(k,l) \tilde{\Omega}_{jl} \mathrm{d} \log p_l
\end{aligned}$$

Note $\sum_k \tilde{\Omega}_{jk} \sum_l \theta_j(k,l) \tilde{\Omega}_{jl} \mathrm{d} \log p_l = \sum_l \tilde{\Omega}_{jl} \mathrm{d} \log p_l \sum_k \tilde{\Omega}_{jk} \theta_j(k,l) = \sum_l \tilde{\Omega}_{jl} \mathrm{d} \log p_l \sum_k \tilde{\Omega}_{jk} \theta_j(l,k)$, where in the last step we use the symmetry of elasticity $\theta_j(l,k) = \theta_j(k,l)$. Further, (F17) implies that this term is zero. Thus

$$\begin{aligned}
\mathrm{d} \log \tilde{\Omega}_{ji} &= \mathrm{d} \log p_i - \sum_k \tilde{\Omega}_{jk} \mathrm{d} \log p_k + \sum_k \theta_j(i,k) \tilde{\Omega}_{jk} \mathrm{d} \log p_k \\
&= \mathrm{d} \log p_i + \sum_k \left(\theta_j(i,k) - 1 \right) \tilde{\Omega}_{jk} \mathrm{d} \log p_k
\end{aligned}$$

Since $\Omega_{ji} = \frac{\tilde{\Omega}_{ji}}{\mu_j}$, we have

$$\mathrm{d}\Omega_{ji} = -\Omega_{ji} \mathrm{d} \log \mu_j + \Omega_{ji} \left(\mathrm{d} \log p_i + \sum_k \left(\theta_j(i,k) - 1 \right) \tilde{\Omega}_{jk} \mathrm{d} \log p_k \right)$$

Using $\Psi = (I - \Omega)^{-1}$ and thus $d\Psi = \Psi d\Omega \Psi$, we get

$$\begin{aligned} d\lambda_n = d\Psi_{0n} &= -\sum_j \sum_i \Psi_{0j}\Omega_{ji} d\log \mu_j \Psi_{in} \\ &+ \sum_j \sum_i \Psi_{0j}\Omega_{ji}\left(d\log p_i + \sum_k \left(\theta_j(i,k) - 1\right)\tilde{\Omega}_{jk} d\log p_k\right)\Psi_{in} \\ &= -\sum_j \Psi_{0j}\left(\Psi_{jn} - 1_{j=n}\right) d\log \mu_j \\ &+ \sum_j \sum_i \sum_l \Psi_{0j}\Omega_{ji}\left(\tilde{\Psi}_{il} + \sum_k \left(\theta_j(i,k) - 1\right)\tilde{\Omega}_{jk}\tilde{\Psi}_{kl}\right) d\log \mu_l \Psi_{in} \end{aligned}$$

where in the last step we plug in (F16). Hence

$$\begin{aligned} \frac{\partial \log \lambda_n}{\partial \log \mu_n} &= -(\Psi_{nn} - 1) + \Psi_{0n}^{-1} \sum_j \sum_i \Psi_{0j}\Omega_{ji}\left(\tilde{\Psi}_{in} + \sum_k \left(\theta_j(i,k) - 1\right)\tilde{\Omega}_{jk}\tilde{\Psi}_{kn}\right)\Psi_{in} \\ &= -(\Psi_{nn} - 1) + \lambda_n^{-1} \sum_j \frac{\lambda_j}{\mu_j}\left[\sum_i \sum_k \tilde{\Omega}_{ji}\left(1_{i=k} + \left(\theta_j(i,k) - 1\right)\tilde{\Omega}_{jk}\right)\tilde{\Psi}_{kn}\Psi_{in}\right] \\ &= -(\Psi_{nn} - 1) - \lambda_n^{-1} \sum_j \frac{\lambda_j}{\mu_j}\Phi_j\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) \end{aligned} \tag{F18}$$

Combining this, (F16) and (F15) yields the results in the proposition.

Last, we show that this proposition nests Proposition 5. When the production function is CES at the industry level, we have $\theta_j(i,k) = \theta_j\left(1 - \frac{1_{i=k}}{\tilde{\Omega}_{ji}}\right)$ and thus

$$\begin{aligned} \Phi_j\left(\tilde{\Psi}_{(l)}, \Psi_{(f)}\right) &= -\sum_{i,k} \tilde{\Omega}_{ji}\left(1_{i=k} + \left(\theta_j\left(1 - \frac{1_{i=k}}{\tilde{\Omega}_{ji}}\right) - 1\right)\tilde{\Omega}_{jk}\right)\tilde{\Psi}_{kl}\Psi_{if} \\ &= -\sum_{i,k} \tilde{\Omega}_{ji}\left(1_{i=k}\left(1 - \theta_j\right) - \left(1 - \theta_j\right)\tilde{\Omega}_{jk}\right)\tilde{\Psi}_{kl}\Psi_{if} \\ &= \left(\theta_j - 1\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\tilde{\Psi}_{(l)}, \Psi_{(f)}\right) \end{aligned}$$

□

Proof of Lemma E1. For two firms $1,2$ in the same industry, with $z_1 > z_2$, take the difference between their Lerner conditions (24) and get

$$\frac{1}{\mu_1 - 1} - \frac{1}{\mu_2 - 1} = -(\rho - \eta)(s_1 - s_2) \tag{F19}$$

Using $\mu_i = \frac{p_i}{\frac{1}{z_i}\mathsf{c}}$ and $s_i = \frac{p_i^{1-\rho}}{p^{1-\rho}}$, we can rewrite the equation in terms of p_1, p_2 as

$$\frac{1}{\frac{z_1 p_1}{\mathsf{c}} - 1} - \frac{1}{\frac{z_2 p_2}{\mathsf{c}} - 1} = -(\rho - \eta)\frac{p_1^{1-\rho} - p_2^{1-\rho}}{p^{1-\rho}}$$

The LHS is decreasing in both z_1 and p_1, and the RHS is increasing in p_1 when $\rho > 1$ and $\rho > \eta$. Suppose by contradiction that $p_1 \geq p_2$, we would have LHS $< 0 \leq$ RHS. Thus for the equality to hold, it must be the case that $p_1 < p_2$. Since s_i is decreasing in p_i, we also have $s_1 > s_2$. Since $\eta < \rho$, from the two firms Lerner condition we conclude that $\mu_1 > \mu_2$.□

Proof of Proposition E1. For simplicity, suppress the industry subscript and denote the more productive firm by 1 and the other firm by 2. Suppose without loss of generality that $z_1 = z > z_2 = 1$. Denote firm 1's market share by s_1 and correspondingly firm 2's market share by $s_2 = 1 - s_1$. Denote the firm elasticity as $e(s) = (1-s)\rho + s\eta$, and the markup is $\mu(s) = \frac{e(s)}{e(s)-1}$. We show that if the industry elasticity η is lower, 1. the more productive firm has a smaller market share (corresponding to lower concentration), 2. both firms have larger markups, 3. the markup-to-market-share relation is steeper, 4. the industry has a lower productivity, and 5. the industry has a higher price index.

1. *Market share.* The market shares satisfy

$$\frac{s_1}{s_2} = \left(\frac{p_1}{p_2}\right)^{1-\rho} = \left(\frac{\mu_1/z}{\mu_2}\right)^{1-\rho} = \left(\frac{\frac{e_1}{e_1-1}\frac{1}{z}}{\frac{e_2}{e_2-1}}\right)^{1-\rho}$$

It can be rewritten as a function of s_1 only,

$$\left(\frac{s_1}{1-s_1}\right)^{\frac{1}{1-\rho}} = \frac{\frac{e(s_1)}{e(s_1)-1}\frac{1}{z}}{\frac{e(1-s_1)}{e(1-s_1)-1}}$$

which implicitly determines the equilibrium s_1 given η. The LHS is decreasing in s_1, approaches ∞ as $s_1 \to 0$, and approaches 0 as $s_1 \to 1$. We want to show that the

equilibrium s_1 increases in η.

Suppose the equilibrium under η features s_1, and we are to find the equilibrium under $\eta' < \eta$. With slight abuse of notation, we evaluate the equation above at s_1 under the lower elasticity η'. If the RHS is larger than the LHS, s_1 is no longer an equilibrium under η'. However, by the intermediate value theorem, there must exist a lower $s_1' < s_1$ that solves the equation. That is, the market share of the more productive firm gets smaller. Now we need to prove that the RHS is larger than the LHS when evaluated at s_1 with η'.

To prove the claim, let e_1, e_2 denote the elasticities of two firms each with market share $s_1, 1 - s_1$ under elasticity η and e_1', e_2' denote the elasticities of two firms still each with market share $s_1, 1 - s_1$ but under the lower elasticity η'. It suffices to show that the RHS under η' is larger than the RHS under η, since the LHS, which is independent of η, is equal to the RHS under η, i.e.,

$$\frac{\frac{e_1'}{e_1'-1}}{\frac{e_2'}{e_2'-1}}\frac{1}{z} > \frac{\frac{e_1}{e_1-1}}{\frac{e_2}{e_2-1}}\frac{1}{z}$$

Rewrite it as

$$\frac{\frac{e_1'}{e_1'-1}}{\frac{e_1}{e_1-1}} - 1 > \frac{\frac{e_2'}{e_2'-1}}{\frac{e_2}{e_2-1}} - 1$$
$$\frac{1-\frac{e_1'}{e_1}}{e_1'-1} > \frac{1-\frac{e_2'}{e_2}}{e_2'-1}$$

For this inequality to hold, it suffices to have $\frac{e_1'}{e_1} < \frac{e_2'}{e_2}$ and $e_1' < e_2'$. Note that the second condition is implied by the first, since $e_1' < \frac{e_1}{e_2}e_2' < e_2'$. Now we prove the first condition, by showing $e_1'e_2 < e_1e_2'$. Plugging in the expression of these elasticities in terms of s_1, we get

$$\begin{aligned}
\left[(1-s_1)\rho + s_1\eta'\right]\left[s_1\rho + (1-s_1)\eta\right] &< \left[(1-s_1)\rho + s_1\eta\right]\left[s_1\rho + (1-s_1)\eta'\right] \\
(1-s_1)^2\rho\eta + s_1^2\rho\eta' &< (1-s_1)^2\rho\eta' + s_1^2\rho\eta \\
(1-2s_1)\rho(\eta - \eta') &< 0
\end{aligned}$$

the last one holds as $s_1 > 1/2, \eta' < \eta$. Since firm 1 is more productive, from Lemma E1 we know that $s_1 > 1/2$, which completes the proof.

2. *Firms' markups.* Now we establish that μ_1, μ_2 both decrease in η. For firm 2, the demand elasticity it faces is $(1 - s_2)\rho + s_2\eta$. Since s_2 decreases in η as shown before, its

demand elasticity increases in η and thus its markup decreases in η. For firm 1, its sales share relative to firm 2 is

$$\frac{s_1}{s_2} = \frac{(\mu_1/z)^{1-\rho}}{\mu_2^{1-\rho}}$$

which gives

$$\mu_1 = z\mu_2 \left(\frac{s_2}{s_1}\right)^{\frac{1}{\rho-1}} \tag{F20}$$

Since μ_2 and s_2/s_1 both decrease in η, we have μ_1 decreases in η.

3. *Markup-to-market-share relation.* Define $\kappa := \frac{\log \mu_1 - \log \mu_2}{\log s_1 - \log s_2}$. Taking log of (F20) yields

$$\log \mu_1 - \log \mu_2 = \log z + \frac{1}{\rho - 1} (\log s_2 - \log s_1)$$

and thus

$$\kappa = \frac{\log z}{\log s_1 - \log s_2} - \frac{1}{\rho - 1}$$

Since we have shown that s_1/s_2 is lower when η is lower, this implies that κ is higher when η is lower.

4. *Industry productivity.* Denote y_1, y_2 as two firms' output quantities, and thus the industry quantity $y_{ind} = \left(y_1^{\frac{\rho-1}{\rho}} + y_2^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}$. Use the expression of industry productivity (19),

$$\frac{1}{z_{ind}} = \frac{1}{z}\frac{y_1}{y_{ind}} + \frac{y_2}{y_{ind}} = \frac{\frac{1}{z}\frac{y_1}{y_2} + 1}{\left[1 + \left(\frac{y_1}{y_2}\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}}$$

Claim that $y_1/y_2 < z^\rho$ and the RHS is decreasing in y_1/y_2 and thus the industry z_{ind} is increasing in y_1/y_2 whenever $y_1/y_2 < z^\rho$. Further, claim that y_1/y_2 is increasing in industry η. Thus, the industry z_{ind} is increasing in industry η.

To prove the second claim, notice that

$$\frac{y_1}{y_2} = \frac{p_1^{-\rho}}{p_2^{-\rho}} = \left(\frac{s_1}{s_2}\right)^{\frac{\rho}{\rho-1}} = \left(\frac{s_1}{1-s_1}\right)^{\frac{\rho}{\rho-1}}$$

which is increasing in s_1. Previously we have shown that s_1 is increasing in η. Thus y_1/y_2 is increasing in η too.

To prove the first claim, notice that

$$\frac{y_1}{y_2} = \frac{p_1^{-\rho}}{p_2^{-\rho}} = \left(\frac{\mu_1/z}{\mu_2}\right)^{-\rho} < z^{\rho}$$

holds as long as $s_1 > s_2$. Define

$$h(x) = \frac{\frac{1}{z}x + 1}{\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}}$$

and it suffices to prove that $h(x)$ is decreasing in x when $x < z^{\rho}$. It suffices to show that $h'(x) < 0$ when $x < z^{\rho}$,

$$\begin{aligned} h'(x) &\propto \frac{1}{z}\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} - \left(\frac{1}{z}x + 1\right)\left(1 + x^{\frac{\rho-1}{\rho}}\right)^{\frac{1}{\rho-1}} x^{-\frac{1}{\rho}} \\ &\propto \frac{1}{z}\left(1 + x^{\frac{\rho-1}{\rho}}\right) - \left(\frac{1}{z}x + 1\right)x^{-\frac{1}{\rho}} \\ &= \frac{1}{z} - x^{-\frac{1}{\rho}} \end{aligned}$$

which is negative as long as $x < z^{\rho}$.

5. *Industry price.* Since the industry price p is a CES aggregate of $p_1 = \mu_1/z$ and $p_2 = \mu_2$ and since both μ_1, μ_2 decrease in η, we have p decreasing in η. □

F.2 Endogenous Markup Responses to Firm-Level Productivity Shocks

Here I derive the linear system (38-40) that determines the endogenous firm markup responses to productivity shocks in Section 6.

Firm block (38). Taking the total differentiation of firm's optimal markup from Proposition 4 yields

$$\begin{aligned} -\frac{\mu_{i_n}}{(\mu_{i_n}-1)^2}\mathrm{d}\log\mu_{i_n} = (\eta_n - \rho)\, s_{i_n}\mathrm{d}\log s_{i_n} + s_{i_n}\mathrm{d}\eta_n + \rho\left(\mathrm{d}\eta_n - \mathrm{d}\eta_n^z\right)\left(\tilde{s}_{i_n} - s_{i_n}\right) \\ + \rho\left(\eta_n - \eta_n^z\right)\left(\tilde{s}_{i_n}\mathrm{d}\log\tilde{s}_{i_n} - s_{i_n}\mathrm{d}\log s_{i_n}\right) \end{aligned} \tag{F21}$$

which describes how each firm's markup μ_{i_n} should change as its own market shares $s_{i_n}, \tilde{s}_{i_n}$ and the industry demand elasticities η_n, η_n^z change. Further, its market shares adjust in response to markup and productivity changes of all firms in the same industry n, as in the proof of Lemma 2,

$$\mathrm{d}\log s_{i_n} = \sum_{j_n \in \mathcal{I}_n} (1-\rho)\left(1_{i_n=j_n} - s_{j_n}\right)\left(\mathrm{d}\log \mu_{j_n} - \mathrm{d}\log z_{j_n}\right) \tag{F22}$$

$$\mathrm{d}\log \tilde{s}_{i_n} = \sum_{j_n \in \mathcal{I}_n} \left(1_{i_n=j_n} - \tilde{s}_{j_n}\right)\left(-\rho \mathrm{d}\log \mu_{j_n} - (1-\rho)\,\mathrm{d}\log z_{j_n}\right) \tag{F23}$$

Jointly, (F21-F23) characterize how firm markups μ_{i_n}, market shares s_{i_n} and $\tilde{s}_{i_n}$ change in response to productivity shocks $\mathrm{d}\log z_{j_n}$ and changes in industry demand elasticities η_n, η_n^z, which we can collect as (38).

Industry block (39). From the proof of Lemma 2, we have

$$\mathrm{d}\log \mu_n = \sum_{i_n} \left[(1-\rho)\, s_{i_n} + \rho \tilde{s}_{i_n}\right] \mathrm{d}\log \mu_{i_n} + (\rho - 1)\left(s_{i_n} - \tilde{s}_{i_n}\right) \mathrm{d}\log z_{i_n} \tag{F24}$$

$$\mathrm{d}\log z_n = \sum_{i_n} -\rho\left(s_{i_n} - \tilde{s}_{i_n}\right) \mathrm{d}\log \mu_{i_n} + \left[\rho s_{i_n} - (\rho - 1)\,\tilde{s}_{i_n}\right] \mathrm{d}\log z_{i_n} \tag{F25}$$

Network block (40). From Proposition 5, we have

$$\mathrm{d}\eta_n = \mathrm{dcycle}_n + \left(\theta_p - 1\right)\mathrm{dSP}_n + (\theta_c - 1)\,\mathrm{dSC}_n + (\epsilon - 1)\,\mathrm{dCP}_n \tag{F26}$$

$$\mathrm{d}\eta_n^z = \mathrm{d}\eta_n - \mathrm{dcycle}_n \tag{F27}$$

Further, since

$$\begin{aligned}
\mathrm{cycle}_n &= \Psi_{nn} - 1 \\
\mathrm{SC}_n &= \frac{1}{\lambda_n} \mathrm{Cov}_{\tilde{\Omega}^{(0)}}\left(\tilde{\Psi}_{(n)}, \Psi_{(n)}\right) = \frac{1}{\lambda_n}\left[\sum_m \tilde{\Omega}_{0m}\tilde{\Psi}_{mn}\Psi_{mn} - \left(\sum_m \tilde{\Omega}_{0m}\tilde{\Psi}_{mn}\right)\left(\sum_m \tilde{\Omega}_{0m}\Psi_{mn}\right)\right] \\
\mathrm{SP}_n &= \tilde{\Psi}_{nn} + \Psi_{nn} - 1 - \tilde{\lambda}_n - \mathrm{SC}_n \\
\mathrm{CP}_n &= \tilde{\lambda}_n
\end{aligned}$$

we have

$$\begin{aligned}
\mathrm{dcycle}_n &= \mathrm{d}\Psi_{nn} \\
\mathrm{dSP}_n &= \mathrm{d}\tilde{\Psi}_{nn} + \mathrm{d}\Psi_{nn} - \mathrm{d}\tilde{\lambda}_n - \mathrm{dSC}_n \\
\mathrm{dCP}_n &= \mathrm{d}\tilde{\lambda}_n \\
\mathrm{d}\lambda_n &= \mathrm{d}\Psi_{0n} \\
\mathrm{d}\tilde{\lambda}_n &= \mathrm{d}\tilde{\Psi}_{0n}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{dSC}_n = &-\frac{\mathrm{d}\lambda_n}{\lambda_n}\mathrm{SC}_n + \frac{\sum_m \left(\mathrm{d}\tilde{\Omega}_{0m}\tilde{\Psi}_{mn}\Psi_{mn} + \tilde{\Omega}_{0m}\mathrm{d}\tilde{\Psi}_{mn}\Psi_{mn} + \tilde{\Omega}_{0m}\tilde{\Psi}_{mn}\mathrm{d}\Psi_{mn}\right)}{\lambda_n} \\
&- \frac{\left(\sum_m \left(\mathrm{d}\tilde{\Omega}_{0m}\tilde{\Psi}_{mn} + \tilde{\Omega}_{0m}\mathrm{d}\tilde{\Psi}_{mn}\right)\right)\left(\sum_m \tilde{\Omega}_{0m}\Psi_{mn}\right) + \left(\sum_m \tilde{\Omega}_{0m}\tilde{\Psi}_{mn}\right)\left(\sum_m \left(\mathrm{d}\tilde{\Omega}_{0m}\Psi_{mn} + \tilde{\Omega}_{0m}\mathrm{d}\Psi_{mn}\right)\right)}{\lambda_n}
\end{aligned}$$

These determine $\mathrm{d}\eta_n, \mathrm{d}\eta_n^z$ in (F26-F27).

To determine the responses of these network statistics to industry markups and productivities, we use

$$\mathrm{d}\tilde{\Omega}_{ji} = \left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\left(\mathrm{d}\log\mu_k - \mathrm{d}\log z_k\right), I_{(i)}\right) \tag{F28}$$

$$\mathrm{d}\Omega_{ji} = -\Omega_{ji}\mathrm{d}\log\mu_j + \frac{1}{\mu_j}\mathrm{d}\tilde{\Omega}_{ji} \tag{F29}$$

As $\tilde{\Psi} = \left(I - \tilde{\Omega}\right)^{-1}, \Psi = (I-\Omega)^{-1}$, we have $\mathrm{d}\tilde{\Psi} = \tilde{\Psi}\mathrm{d}\tilde{\Omega}\tilde{\Psi}, \mathrm{d}\Psi = \Psi\mathrm{d}\Omega\Psi$ and thus

$$\mathrm{d}\tilde{\Psi}_{mn} = \sum_j \tilde{\Psi}_{mj}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\left(\mathrm{d}\log\mu_k - \mathrm{d}\log z_k\right), \tilde{\Psi}_{(n)}\right) \tag{F30}$$

$$\begin{aligned}
\mathrm{d}\Psi_{mn} &= -\sum_j \Psi_{mj}\mathrm{d}\log\mu_j \sum_i \Omega_{ji}\Psi_{in} + \sum_j \frac{\Psi_{mj}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\left(\mathrm{d}\log\mu_k - \mathrm{d}\log z_k\right), \Psi_{(n)}\right) \\
&= -\sum_j \Psi_{mj}\mathrm{d}\log\mu_j\left(\Psi_{jn} - 1_{j=n}\right) + \sum_j \frac{\Psi_{mj}}{\mu_j}\left(1-\theta_j\right)\mathrm{Cov}_{\tilde{\Omega}^{(j)}}\left(\sum_k \tilde{\Psi}_{(k)}\left(\mathrm{d}\log\mu_k - \mathrm{d}\log z_k\right), \Psi_{(n)}\right)
\end{aligned} \tag{F31}$$

Together, these form the linear system that maps changes in industry markups and productivities to industry elasticities.

Bibliography

Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry. 2012. "Measuring the Upstreamness of Production and Trade Flows." *American Economic Review: Papers & Proceedings*, 102(3): 412–416. A-3

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review*, 98(5): 1998–2031. A-9

Autor, David H., David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *The Quarterly Journal of Economics*, 135(2): 645–709. A-12

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and Misallocation in General Equilibrium." *The Quarterly Journal of Economics*, 135(1): 105–163. A-2, A-10, A-22

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. "The Rise of Market Power and the Macroeconomic Implications." *The Quarterly Journal of Economics*, 135(2): 561–644. A-12

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. 2026. "The Hitchhiker's Guide to Markup Estimation: Assessing Estimates From Financial Data." *Econometrica*, 94(1): 137–168. A-2

Gutiérrez, Germán, and Thomas Philippon. 2017. "Investmentless Growth: An Empirical Investigation." *Brookings Papers on Economic Activity*, 2017(2): 89–190. A-12

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies*, 75(1): 295–316. A-14